

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
8 November 2001 (08.11.2001)

PCT

(10) International Publication Number
WO 01/83746 A2

(51) International Patent Classification⁷: **C12N 15/12**,
C07K 14/705, C12Q 1/68, A61K 31/7088, A01K 67/027

[FR/FR]; 4, rue Rollin, F-75005 Paris (FR). **REMALEY, Alan** [US/US]; 4510 Traymore Street, Bethesda, MD 20892 (US). **SANTAMARINA-FOJO, Silvia** [US/US]; 10805 Pebble Brook Lane, Potomac, MD 20854 (US).

(21) International Application Number: PCT/EP01/05488

(22) International Filing Date: 2 May 2001 (02.05.2001)

(74) Agent: **LECCA, Patricia**; Aventis Pharma S.A., Direction Brevets, 20, avenue Raymond Aron, F-92165 Antony Cedex (FR).

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/201,280 2 May 2000 (02.05.2000) US

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(71) Applicant (*for all designated States except US*): **AVENTIS PHARMA S.A.** [FR/FR]; 20, avenue Raymond Aron, F-92160 Antony (FR).

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): **ROSIER-MONTUS, Marie-Françoise** [FR/FR]; 21, rue des Baconnets, F-92160 Antony (FR). **PRADES, Catherine** [FR/FR]; 30, avenue du Général De Gaulle, F-94320 Thiais (FR). **LEMOINE, Cendrine** [FR/FR]; 5, avenue Marcel Ramolfo Garnier, F-91300 Massy (FR). **NAUDIN, Laurent** [FR/FR]; 11 bis, rue de la Roche Plate, F-91150 Etampes (FR). **DENEFLÉ, Patrice** [FR/TR]; 45, avenue des Fusillés de Chateaubriand, F-94160 Saint-Maur (FR). **BREWER, Bryan** [US/US]; 10805 Pebble Brook Lane, Potomac, MD 20854 (US). **DUVERGER, Nicolas**

Published:

— without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.



WO 01/83746 A2

(54) Title: REGULATORY NUCLEIC ACID FOR THE ABC1 GENE, MOLECULES MODIFYING ITS ACTIVITY AND THERAPEUTIC USES

(57) Abstract: The present invention concerns a nucleic acid which is capable of regulating the transcription of the ABC1 gene, which is a causal gene for pathologies linked to a dysfunctioning of cholesterol metabolism, inducing diseases such as atherosclerosis. The invention also relates to nucleotide constructs comprising a polynucleotide which encodes a polypeptide or a nucleic acid of interest, placed under the control a regulatory nucleic acid for the ABC1 gene. The invention also relates to recombinant vectors, transformed host cells and nonhuman transgenic mammals comprising a nucleic acid which regulates the transcription of the ABC1 gene or an abovementioned nucleotide construct, as well as methods for screening molecules or substances which are capable of modifying the activity of the regulatory nucleic acid for the ABC1 gene.

**REGULATORY NUCLEIC ACID FOR THE ABC1 GENE,
MOLECULES MODIFYING ITS ACTIVITY AND
THERAPEUTIC USES**

The present invention concerns a nucleic acid which is capable of regulating the transcription of the ABC1 gene, which is a causal gene for pathologies linked to a dysfunctioning of cholesterol metabolism, inducing diseases such as
5 atherosclerosis.

The invention also relates to nucleotide constructs comprising a polynucleotide which encodes a polypeptide or a nucleic acid of interest, placed under the control of a regulatory nucleic acid for the ABC1 gene.

The invention also relates to recombinant vectors, transformed host cells and
10 non-human transgenic mammals comprising a nucleic acid which regulates the transcription of the ABC1 gene or an above-mentioned nucleotide construct, as well as to methods for screening molecules or substances which are capable of modifying the activity of the regulatory nucleic acid for the ABC1 gene.

The invention also relates to methods for detecting an impairment in the
15 transcription of the ABC1 gene in an at-risk individual.

A subject of the invention is also substances or molecules which modify the activity of the nucleic acid which regulates the transcription of the ABC1 gene, as well as pharmaceutical compositions containing such substances or such molecules.

20 High density lipoproteins (HDLs) are one of the four major classes of lipoprotein which circulate in the blood plasma.

These lipoproteins are involved in various metabolic pathways, such as lipid transport, bile acid formation, steroidogenesis or cell proliferation, and also interfere with plasmatic proteinase systems.

HDLs are perfect free cholesterol acceptors, and in combination with
5 cholesterol ester transfer proteins (CETP), lipoprotein lipase (LPL), hepatic lipase (HL) and lecithin-cholesterol acyltransferase (LCAT), play a major role in the reverse transport of cholesterol, i.e. the transport of excess cholesterol in the peripheral cells to the liver, for its removal from the body in the form of bile acid.

It has been demonstrated that the HDLs generally play a central role in the
10 transport of cholesterol from the peripheral tissues to the liver.

Various diseases linked to an HDL deficiency have been described, including Tangier disease, HDL deficiency and LCAT deficiency.

The deficiency involved in Tangier disease is linked to a cellular defect in the translocation of cellular cholesterol, which leads to a degradation of the HDLs.

15 In Tangier disease, this cellular defect leads to a disruption of lipoprotein metabolism. The HDL particles in Tangier disease, which do not incorporate cholesterol from the peripheral cells, and which are not able to be correctly metabolized, are rapidly eliminated from the body. The plasma HDL concentration in these patients is thus extremely reduced, and the HDLs no longer contribute to
20 the return of cholesterol to the liver. This cholesterol accumulates in these peripheral cells and causes characteristic clinical manifestations such as the formation of orange-colored tonsils. Furthermore, other lipoprotein disruptions such

as an overproduction of triglycerides and an increased synthesis and intracellular catabolism of phospholipids are generally observed.

Tangier disease, the symptoms of which have been described above, is classified among the familial conditions linked to metabolism of the HDLs which are
5 commonly detected in patients affected by coronary diseases.

Numerous studies have shown that a reduced level of HDL cholesterol is a risk factor which is useful for detecting a coronary condition.

In this context, syndromes linked to HDL deficiencies have been of increasing interest for the past decade, since they make it possible to increase the
10 understanding of the role of HDLs in atherogenesis.

Several mutations in the apo A-1 gene have been characterized. These mutations are rare and can lead to an absence of production of apo A-1.

Mutation in the genes encoding lipoprotein lipase (LPL) or its activator apoC-II are associated with severe hypertriglyceridemias and substantially reduced
15 levels of HDL-c.

Mutations in the gene encoding the enzyme lecithin-cholesterol acyltransferase (LCAT) are also associated with severe HDL deficiency.

Furthermore, dysfunctions in the reverse transport of cholesterol might be induced by physiological deficiencies affecting at least one of the steps for
20 transporting stored cholesterol from the intracellular vesicles toward the membrane surface, where it is accepted by the HDLs.

Recently, a study was carried out on the segregation of various allelic forms of 343 microsatellite markers distributed over the entire genome and distant from each other by 10.3 cM on average.

The linkage study was carried out on a family which had been well
5 characterized over eleven generations, in which many members are affected by Tangier disease, the family comprising five consanguineous lines.

This study made it possible to identify a region located in the 9q31 locus of human chromosome 9 which is statistically linked to the condition (Rust S. et al., Nature Genetics Vol. 20, September 1998, pages 96-98).

10 However, the study by Rust et al. only characterizes a wide region of the genome in which impairments are likely to be associated with Tangier disease. The study simply stated that the relevant 9q31-34 region contains ESTs, but no known gene.

It has been shown that a region spanning 1 cM, situated in the 9q31 locus in
15 humans, is generally associated with familial HDL deficiencies (Rust et al., 1999).

Furthermore, it has been shown that a gene encoding a protein of the family of ABC transporters, which is located precisely in the 1 cM region of the 9q31 locus, is involved in pathologies linked to a deficiency in the reverse transport of cholesterol.

20 For example, it has been shown that the gene encoding the ABC-1 transporter is mutated in patients with affected reverse transport of cholesterol, such as in patients suffering from Tangier disease.

The ABC ("ATP-binding cassette") transporter proteins constitute a family of proteins which are extremely conserved in evolution, from bacteria to humans.

The ABC transporter proteins are involved in the membrane transport of various substrates, for example, ions, amino acids, peptides, sugars, vitamins or
5 steroid hormones.

The characterization of the complete amino acid sequence of some ABC transporters has made it possible to determine that these proteins have a common general structure, for example, two nucleotide-binding folds (Nucleotide Binding Fold or NBF) with moieties of Walker A type and Walker B type, as well as two
10 transmembrane domains, each of the transmembrane domains consisting of six helices. The specificity of the ABC transporters for the various transported molecules appears to be determined by the structure of the transmembrane domains, whereas the energy required for the transport activity is provided by degrading ATP at the NBF fold.

15 Several of the ABC transporter proteins which have been identified in humans have been associated with various diseases.

For example, cystic fibrosis is caused by mutations in the CFTR (cystic fibrosis transmembrane conductance regulator) gene.

Moreover, some multi-drug resistance phenotypes in tumor cells have been
20 associated with mutations in the gene encoding the MDR (multi-drug resistance) protein which also has an ABC transporter structure.

Other ABC transporters have been associated with neuronal and tumor conditions (US patent No. 5,858,719), or potentially implicated in diseases caused by an impairment of the homeostasis of metals, for example, the ABC-3 protein.

Similarly, another ABC transporter, referred to as PFIC2, seems to be
5 involved in a form of progressive familial intrahepatic cholestasis, this protein being potentially responsible, in humans, for the export of bile salts.

In 1994, a cDNA encoding a novel mouse ABC transporter was identified and referred to as ABC1 (Luciani et al., 1994). This protein is characteristic of the ABC transporters in that it has a symmetrical structure comprising two transmembrane
10 domains linked to a highly hydrophobic segment and to two NBF moieties.

In humans, a partial cDNA comprising the entire open reading frame of the human ABC1 transporter has been identified (Langmann et al., 1999).

It has also been shown that the gene encoding the human ABC1 protein is expressed in various tissues, and more particularly at high levels in the placenta, the
15 liver, the lungs, the adrenal glands and the fetal tissues.

These authors have also shown that the expression of the gene encoding the human ABC1 protein is induced during the differentiation of monocytes into macrophages *in vitro*. Furthermore, the expression of the gene encoding the ABC1 protein is increased when human macrophages are incubated in the presence of
20 acetylated low-density lipoproteins (AcLDLs).

The work of Rust S. et al., 1999, Brooks-Wilson A. et al., 1999, Bodzioch M. et al., 1999, Remaley A. et al., 1999 and of Marcil M. et al., 1999 has shown that patients suffering from Tangier disease and from HDL deficiencies (FHD; familial

HDL deficiency) have a mutated ABC1 gene. Several mutations, which are distributed in various regions of the ABC1 gene, have been identified in the genome of various patients, for example, of patients with a severe form of the disease associated with coronary disorders. Moreover, diverse polymorphisms have been
5 found, both in the exons and in the introns of the ABC1 gene, in patients suffering from milder forms of the disease, indicating that these patients carry specific alleles of the gene, which are distinct from the "wild-type" allele(s).

Although the expression of the human ABC1 gene seems to be regulated according to the type of cell or to the metabolic situation of a given cell type, the
10 sequence(s) which make(s) it possible to regulate this gene were not known.

Thus, there exists a need in the state of the art to identify these regulatory sequences, for the two principal reasons below:

a) These sequences are likely to be mutated in patients suffering from a pathology linked to a deficiency in cholesterol transport,, for example, in patients
15 suffering from Tangier disease, or likely to develop such pathologies.

The characterization of the regulatory sequences of the human ABC1 gene would make it possible, firstly, to detect mutations in patients, and, for example, also to diagnose the individuals who belong to at-risk familial groups. In addition, the isolation of these regulatory sequences would make it possible to complement the
20 mutated sequence with a functional sequence capable of overcoming the metabolic dysfunctions induced by the mutation(s) diagnosed, through the construction of targeted therapeutic means, such as means intended for gene therapy.

b) The characterization of the regulatory sequences of the ABC1 gene would place at the disposal of persons skilled in the art means capable of allowing the construction, by genetic engineering, and then the expression, of given genes in the cell types in which the ABC1 gene is expressed.

5 c) Moreover, some portions of the regulatory sequences of the ABC1 gene might constitute high expression-level constitutive promoter sequences, which are liable to enable the construction of novel means for expressing given sequences in cells, completing an already existing set of means.

To date, despite the efforts undertaken, the regulatory sequences of the
10 ABC1 gene have remained totally unknown.

The inventors have henceforth isolated and then sequenced a genomic DNA comprising the first two exons of the ABC1 gene (respectively exon 1A and exon 1B), as well as a non-transcribed region of approximately 2.9 kb, which is located on the 5' side of exon 1A, and which comprises regulation signals for the ABC1 gene.

15

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1: illustrates a portion of the sequence SEQ ID No. 3, which starts with the nucleotide at position 1 of the sequence SEQ ID No. 3. The position of each of
20 the characteristic binding moieties for various transcription factors is represented by boxes, the designation of the transcription vector specific for the corresponding sequence being indicated above the nucleotide sequence.

Figure 2A: is a schematic illustration that shows the location of the point mutations introduced into the -200 bp promoter region of the human ABC1 gene;

Figure 2B: displays the luciferase activity in RAW cells after transfection with wildtype and mutant constructs. The data shown represents the mean of 3 independent transfection studies. Values are expressed relative to the wildtype construct;

Figure 3: RAW 264.7 cells were transfected with wild type construct (p200-L) or mutant LXR construct (p200-LXRm) along with a β -galactosidase expression plasmid. Three hours after transfection, cells were refed with fresh media containing 10% FCS. Sixteen hours later, cells were washed with PBS and replaced with DMEM media containing 0.1% BSA and added 50 μ g/ml cholesterol, 2 μ g/ml 22(R)-HOch, 10 μ M 9CRA or , 2 μ g/ml 22(R)-HOch plus 10 μ M 9CRA for 24 hours. Cell lysates were analyzed for luciferase and β -galactosidase activity. Luciferase values were normalized to β -galactosidase and expressed as mean \pm SEM.

Figure 4: RAW 264.7 cells were transfected with wild type construct (p200-L), mutant E-box construct (p200-EBm) or the deleted E-box construct (p200-EBd) along with a β -galactosidase expression plasmid. Addition of 50 μ g/ml cholesterol, 2 μ g/ml 22(R)-HOch, 10 μ M 9CRA or , 2 μ g/ml 22(R)-HOch plus 10 μ M 9CRA were performed. Cell lysates were analyzed for luciferase and β -galactosidase activity.

Luciferase values were normalized to β -galactosidase activity and expressed as mean \pm SEM.

Figure 5: A radiolabelled fragment (-171 to -77 bp) of the human ABC1 promoter was digested with different concentrations of DNaseI in the presence (+) or absence (-) of RAW cell nuclear extract (NE). The G and G+A ladders from Maxam and Gilbert sequence of the radiolabelled fragment are shown. The nucleotide position relative to the transcriptional start site and location of the Sp1, AP1 and EB motifs in the hABC1 promoter are indicated on the right;

Figure 6A: illustrates the probes used for the gel shift analysis. Fragment A (100 bp) includes binding motifs for Sp1 and AP1 and the E-box. Fragment EB (27 bp) contains the E-box and fragment EBm (not shown) contains a mutated E-box. In panels B and C, the labelled fragments (fragments A, EB or EBm) used for the gel shift study are shown on top of each gel. Incubation of the radiolabelled probe with RAW cell nuclear extract (NE) is indicated (+);

Figure 6B: shows the gel-shift analysis performed by incubating RAW cell nuclear extracts with radiolabelled Fragment A (left), EB (middle) or EBm (right) in the presence or absence of specific competitors (unlabelled fragment A, EB or EBm);

Figure 6C: shows supershift analysis of Fragments A or EB with antibodies specific to the amino (N) or carboxyl (C) ends of USF1 and USF2. Arrows indicate

the position of probe complexed with protein and arrows with an asterisk indicates the position of antibody-supershifted complex.

DETAILED DESCRIPTION OF THE INVENTION

5

GENERAL DEFINITIONS

For the purpose of the present invention, the term "isolated" refers to a biological material (nucleic acid or protein) which has been abstracted from its environment of origin (the environment in which it is naturally located).

10 For example, a polynucleotide present in the natural state in a plant or an animal is not isolated. The same polynucleotide separated from the adjacent nucleic acids among which it is naturally inserted in the genome of the plant or the animal is considered to be "isolated".

Such a polynucleotide can be included in a vector and/or such a
15 polynucleotide can be included in a composition, and can remain, however, in the isolated state, because the vector or the composition does not constitute its natural environment.

The term "purified" does not require the material to be present in a form of absolute purity, excluding the presence of other compounds. It is, rather, a relative
20 definition.

A polynucleotide is in the "purified" state after purification of the starting material, or after purification of the natural material, wherein impurities are present after the purification in an amount of at least one order of magnitude less than the

amount of impurities before the purification. In one embodiment, impurities are present after the purification in an amount of 2 or 3 orders of magnitude less than the amount of impurities before the purification. In another embodiment, impurities are present after the purification in an amount of 4 or 5 orders of magnitude less than the amount of impurities before the purification.

For the purposes of the present description, the expression "nucleotide sequence" can be used to refer indiscriminately to a polynucleotide or a nucleic acid.

The expression "nucleotide sequence" encompasses the genetic material itself, and is thus not restricted to the information concerning its sequence.

10 The terms "nucleic acid", "polynucleotide", "oligonucleotide" and "nucleotide sequence" encompass RNA, DNA or cDNA sequences, or DNA/RNA hybrid sequences of more than one nucleotide, indiscriminately in the single chain form or in the duplex form.

The term "nucleotide" refers to both natural nucleotides (A, T, G, C) and modified nucleotides, which comprise at least one modification such as (1) a purine analog, (2) a pyrimidine analog or (3) a similar sugar, examples of such modified nucleotides being described for example, in PCT application No. WO 95/04 064.

For the purposes of the present invention, a first polynucleotide is considered as being "complementary" to a second polynucleotide when each base of the first nucleotide is paired with the complementary base of the second polynucleotide, the orientation of which is inversed. The complementary bases are A and T (or A and U), or C and G.

"Variant" of a nucleic acid according to the invention will be intended to mean a nucleic acid which differs by one or more bases with respect to the reference polynucleotide. A variant nucleic acid can be of natural origin, such as an allelic variant found naturally, or can also be an unnatural variant obtained for example, by
5 mutagenesis techniques.

In general, the differences between the reference nucleic acid and the variant nucleic acid are small, such that the nucleotide sequences of the reference nucleic acid and of the variant nucleic acid are very close and, in many regions, identical. The nucleotide modifications present in a variant nucleic acid can be silent, which
10 means that they do not alter the amino acid sequences encoded by said variant nucleic acid.

However, the nucleotide changes in a variant nucleic acid can also result in substitutions, additions or deletions in the polypeptide encoded by the variant nucleic acid, with respect to the peptides encoded by the reference nucleic acid. In
15 addition, nucleotide modifications in the coding regions can produce substitutions, which may be conservative or non-conservative in the amino acid sequence.

For example, the variant nucleic acids according to the invention encode polypeptides which conserve more or less the same function or biological activity as the polypeptide of the reference nucleic acid, or the capacity to be recognized by
20 antibodies directed against the polypeptides encoded by the initial nucleic acid.

Some variant nucleic acids will thus encode mutated forms of the polypeptides, the systematic study of which will make it possible to deduce structure-activity relationships for the proteins in question. Knowledge of these

variants with respect to the disease studied is fundamental, since it makes it possible to understand the molecular cause of the pathology.

A "fragment" of a reference nucleic acid according to the invention, will be intended to mean a nucleotide sequence which is shorter in length than the nucleotide sequence of the reference nucleic acid and which comprises a nucleotide sequence which is identical to a portion of the nucleotide sequence of the reference nucleic acid. Such a "fragment" of nucleic acid according to the invention can be, if needed, included in the nucleotide sequence of a second polynucleotide different from the reference nucleic acid. The resulting nucleotide which comprises the "fragment" and the second polynucleotide may have a nucleotide sequence that is longer than, the same length as, or shorter than the nucleotide sequence of the reference nucleic acid. Such fragments comprise, or alternatively consist of, oligonucleotides of lengths ranging from 20 to 25, 30, 40, 50, 70, 80, 100, 200, 500, 1000 or 1500 consecutive nucleotides of a nucleic acid according to the invention.

"Biologically active fragment" of an acid which regulates transcription according to the invention is intended to mean a nucleic acid which is capable of modifying the transcription of a sequence of DNA placed under its control. Such a biologically active fragment comprises a basic promoter and/or a regulatory element, as defined in the present description.

"Regulatory nucleic acid" according to the invention is intended to mean a nucleic acid which activates and/or regulates the expression of a DNA sequence which is selected and placed under its control.

"Promoter" is intended to mean a DNA sequence recognized by the proteins of the cell which are involved in initiating the transcription of a gene. The basic promoter is the minimum regulatory nucleic acid which is capable of initiating the transcription of a given DNA sequence which is placed under its control. In general, the basic promoter consists a region of genomic DNA upstream of the transcription start site, where a sequence CAAT (to which one or more protein transcription factors bind), as well as, except in rare cases such as in certain housekeeping genes, the sequence TATA, or "TATA box", or a related box, are very often found. An RNA polymerase, as well as one or more transcription factors, such as "TATA" box binding proteins (or TBPs), bind to this box.

A nucleotide sequence is "placed under the control" of a regulatory nucleic acid when this regulatory nucleic acid is located, with respect to the nucleotide sequence, in such a way as to control the initiation of the transcription of the nucleotide sequence with an RNA polymerase.

For the purpose of the invention, "regulatory element" and "regulatory sequence" are intended to mean a nucleic acid comprising elements capable of modifying the transcription initiated by a basic promoter, such as binding sites for diverse transcription factors, "enhancer" sequences for increasing transcription or "silencer" sequences for inhibiting transcription.

"Enhancer" sequence is intended to mean a DNA sequence included in a regulatory nucleic acid, which is capable of increasing or of stimulating the transcription initiated by a basic promoter.

"Silencer" sequence is intended to mean a DNA sequence included in a regulatory acid, which is capable of decreasing or of inhibiting the transcription initiated by a basic a promoter.

Regulatory elements can be present outside the sequence which is located
5 on the 5' side of the transcription start site, for example, in the introns and the exons, including in the coding sequences.

The terms "basic promoter" and "regulatory element" can be "specific for one or more tissues" if they allow the transcription of a given DNA sequence, placed under their control, such as in certain cells (for example, the cells specific for a
10 tissue), i.e. either exclusively in the cells of certain tissues, or at different levels of transcription according to the tissues.

"Transcription factor" is intended to mean proteins which preferentially interact with regulatory elements of a regulatory nucleic acid according to the invention, and which stimulate or, on the contrary, suppress transcription. Some
15 transcription factors are active in the form of monomers, others being active in the form of homo- or heterodimers.

The term "modification" is directed toward either a positive regulation (increase, stimulation) of transcription, or a negative regulation (decrease, inhibition, blocking) of transcription.

20 For the purpose of the present invention, the "percentage of identity" between two sequences of nucleotides or amino acids can be determined by comparing two optimally aligned sequences, through a comparison window.

The portion of the nucleotide sequence or polypeptide which is in the comparison window can thus comprise additions or deletions (for example, "gaps"), with respect to the reference sequence (which does not comprises these additions or these deletions), in such a way as to obtain an optimal alignment of the two
5 sequences.

The percentage is calculated by determining the number of positions at which an identical nucleic base or amino acid residue is observed for the two sequences (nucleic or peptide) compared, then by dividing the number of positions at which there is identity between the two bases or amino acid residues by the total number
10 of positions in the comparison window, and then by multiplying the result by 100 to obtain the percentage of sequence identity.

The optimal sequence alignment for the comparison can be carried out using a computer with the aid of known algorithms in the package from the company Wisconsin Genetics Software Package, Genetics Computer Groups (GCG), 575
15 Science Doctor, Madison, Wisconsin.

By way of illustration, it will be possible to produce the percentage of sequence identity with the aid of the BLAST software (Versions BLAST 1.4.9 of March 1996, BLAST 2.0.4 of February 1998 and BLAST 2.0.6 of September 1998), using exclusively the default parameters (S. F. Altschul et al., J. Mol. Biol. 1990 215:
20 403-410, S. F. Altschul et al., Nucleic Acids Res. 1997 25: 3389-3402). Blast searches for sequences which are similar/homologous to a "request" sequence of reference, with the aid of the algorithm of Altschul et al. The request sequence and

the databases used can be of peptide or nucleic type, any combination being possible.

For the purposes of the present invention, "high stringency hybridization conditions" are intended to mean the following conditions:

5

1 - Membrane competition and PREHYBRIDIZATION:

- Mixed: 40 μ l salmon sperm DNA (10 mg/ml)

+ 40 μ l human placental DNA (10 mg/ml)

- Denatured for 5 min at 96°C, then immersed the mixture in ice.

10 - Removed the 2X SSC buffer and poured 4 ml of formamide mix into the hybridization tube containing the membranes.

- Added the mixture of the two denatured DNAs.

- Incubation at 42°C for 5 to 6 hours, with rotation.

15 **2 - Labeled probe competition:**

- Added 10 to 50 μ l Cot I DNA to the labeled and purified probe, according to the amount of nonspecific hybridization.

- Denatured for 7 to 10 min at 95°C.

- Incubated at 65°C for 2 to 5 hours.

20

3 - HYBRIDIZATION:

- Removed the prehybridization mix.

- Mixed 40 μ l salmon sperm DNA + 40 μ l human placental DNA; denatured 5 min at 96°C, then immersed in ice.
- Added 4 ml of formamide mix, the mixture of the two DNAs and the denatured labeled probe/Cot I DNA to the hybridization tube.
- 5 - Incubated for 15 to 20 hours at 42°C, with rotation.

4 - Washes:

- One wash at room temperature in 2X SSC, to rinse.
- Twice 5 minutes at room temperature, 2X SSC and 0.1% SDS, at 65°C.
- 10 - Twice 15 minutes at 65°C, 1X SSC and 0.1% SDS, at 65°C.

Wrapped the membranes in Saranwrap and exposed.

The hybridization conditions described above are adapted to hybridization, under high stringency conditions, of a molecule of nucleic acid of variable length, from 20 nucleotides to several hundred nucleotides.

The hybridization conditions described above can be adapted as a function of the length of the nucleic acid whose hybridization is desired, or of the type of labeling chosen, according to techniques known to persons skilled in the art.

Suitable hybridization conditions may for example, be adapted according to the teaching contained in the work by Hames and Higgins (1985) or in the work by F. Ausubel et al. (1999).

For the purposes of the present invention, "transformation" is intended to mean the introduction of a nucleic acid (or of a recombinant vector) into a host cell.

The term "transformation also encompasses a situation in which the genotype of a cell has been modified by an exogenous nucleic acid, and in which this cell thus transformed expresses said exogenous nucleic acid, for example, in the form of a recombinant polypeptide or in the form of a sense or antisense nucleic acid.

5 For the purposes of the invention, "transgenic animal" is intended to mean a nonhuman animal, such as a mammal, in which one or more cells contain a heterologous nucleic acid which has been introduced through human intervention, such as by transgenesis techniques well known to persons skilled in the art. The heterologous nucleic acid is introduced directly or indirectly into the cell or the
10 precursor of the cell, by genetic manipulation such as microinjection or infection with a recombinant virus. The heterologous nucleic acid can be integrated into the chromosome, or can be in the form of DNA which replicates extrachromosomally.

REGULATORY NUCLEIC ACID FOR THE ABC1 GENE

15 The inventors have succeeded in isolating a regulatory nucleic acid for the human ABC1 gene from vector libraries of BAC type prepared from human genomic material.

According to the sequence analysis carried out, the inventors have determined that the nucleic acid which regulates the transcription of the ABC1 gene,
20 when it is defined in the broadest way, consists of a polynucleotide comprising, from the 5' end toward the 3' end:

- a nontranscribed region of approximately 2.9 kb located upstream of the transcription start site of the ABC1 gene;

- the nucleotide sequence of the first exon of the ABC1 gene, also referred to as exon 1A;

- the partial nucleotide sequence of the first intron of the ABC1 gene, also referred to under the name intron 1A;

5 and

- the nucleotide sequence of the second exon of the human ABC1 gene, also referred to as exon 1B.

- the partial nucleotide sequence of the second intron of the ABC1 gene, also referred under the name intron 1B;

10 Under a general definition, the nucleic acid which regulates the transcription of the ABC1 gene comprises all the nucleotide regions as defined above, and is identified as the sequence SEQ ID No. 1 according to the invention.

Preferentially, the nucleic acid which regulates the transcription factor of the ABC1 gene comprises all the nucleotide regions comprising nucleotide -2228 to
15 nucleotide +108 with respect of the transcription start site of the ABC1 gene, i.e., a region comprised in nucleotides 654 to 3001 of sequence SEQ ID NO: 1.

Thus, a first subject of the invention consists of a nucleic acid comprising a polynucleotide which has at least 20 consecutive nucleotides of the nucleotide sequence SEQ ID No. 1, or a nucleic acid of complementary sequence.

20 The region of approximately 2.9 kb, which is located upstream of the transcription start site of the ABC1 gene, and which comprises the basic promoter and multiple regulatory elements for transcription, is also included in the sequence identified as SEQ ID No. 3 according to the invention.

More precisely, the nucleotide in position 1 of the sequence SEQ ID No. 3 is the nucleotide in position -2893, with respect to the transcription start site of the ABC1 gene.

According to a second aspect, the invention relates to a nucleic acid
5 comprising a polynucleotide which has at least 20 consecutive nucleotides of the nucleotide sequence SEQ ID No. 3, or to a nucleic acid of complementary sequence.

As already specified above, in addition to a 5' non transcribed regulatory region, the nucleic acid which regulates the transcription of the ABC1 gene, of
10 sequence SEQ ID No. 1, also comprises the first exon and the 5' portion of the first intron of the human ABC1 gene.

The first exon of the ABC1 gene, also referred to as exon 1A, is defined as the sequence SEQ ID No. 4.

The sequence of the intron 1a has been partially characterized. The 5' end
15 of intron 1a is defined as the nucleotide sequence SEQ ID No. 6. The 3' end of intron 1a is defined as the sequence SEQ ID No. 7.

The second exon of the human ABC1 gene, also referred to as exon 1B, is defined as the sequence SEQ ID No. 5.

According to a third aspect, the invention relates to a nucleic acid comprising
20 a polynucleotide which has at least 20 consecutive nucleotides of a nucleotide sequence chosen from the sequences SEQ ID No. 3 to 7, or to a nucleic acid of complementary sequence.

In some embodiments of the invention, a nucleic acid according to the invention will be in an isolated and/or purified form.

Any "biologically active" fragment of a nucleic acid as defined above also forms part of the invention.

5 According to another aspect, the invention concerns a nucleic acid having at least 80% nucleotide identity with a nucleic acid as defined above.

The invention also encompasses a nucleic acid that hybridizes, under high stringency conditions, with any one of the nucleic acids according to the invention.

The invention also concerns a nucleic acid having at least 80%, for example,
10 90%, 95% or 98%, nucleotide identity with a nucleic acid comprising at least 20 consecutive nucleotides of a polynucleotide chosen from nucleotide sequences SEQ ID No. 1, SEQ ID No. 2, SEQ ID No. 3, SEQ ID No. 4, SEQ ID No. 5, SEQ ID No. 6, SEQ ID No. 7, and SEQ ID No. 8.

DETAILED ANALYSIS OF THE SEQUENCE SEQ ID NO. 3

15 The nucleic acid of sequence SEQ ID No. 3, included in the regulatory nucleic acid for the ABC1 gene of sequence SEQ ID No. 1, comprises the constituent elements of a basic promoter, respectively a "TATA" box and a homeobox, represented in Figure 1.

The regulatory sequence SEQ ID No. 3 also comprises many binding sites for
20 diverse transcription factors which are capable of positively or negatively regulating the activity of the basic promoter.

Thus, the various sequences which are characteristic of the binding sites for diverse transcription factors in the sequence SEQ ID No. 3 have been identified by the inventors, in the manner detailed below.

The sequence SEQ ID No. 3 was used as a reference sequence and
5 processed according to the algorithms of the BLAST 2, version 2.10, software, and compared to the data listed in several databases, and the presence, as well as the location, of the various characteristic sites of the sequence SEQ ID No. 3, and, for example, the binding sites for transcription factors, were determined according to methods well known to persons skilled in the art.

10 Furthermore, a detailed analysis was carried out on the 1.3 kb upstream of the start site, in which a total of 1900 sequences corresponding to binding sites for transcription factors was identified during the first step of the search. After compilation and filtering as described above, only 79 binding sites, specific for 27 different transcription factors, were retained. These sites are presented in Table 1
15 below.

Table 1 represents the binding sites for the transcription factors identified in the 1318 nt in the 3' portion of the sequence SEQ ID No. 3 according to the invention.

The positions of the start and end nucleotides of each of the binding sites for
20 transcription factors are referred with reference to the numbering of the nucleotides of the sequence SEQ ID No. 3, as represented in Figure 1 (+ strand) or of the nucleotides on the complementary sequence of the sequence SEQ ID No. 3 (- strand)

TABLE 1

Transcription factor		Start	End	Strand
RAR		-1299	-1294	+
TFIID		-1251	-1246	+
AP2		-1203	-1192	+
MYB		-1177	-1170	+
GR		-1166	-1153	+
T3R		-1110	-1095	-
AP1 (C-JUN)		-1052	-1046	+
CEBP(CEBPA CEBPB)		-1015	-1008	+
SP1		-992	-986	+
HNF(HNF3, HNF5)		-951	-940	-
MYB		-941	-934	+
T3R		-933	-928	+
CP2		-924	-918	-
TFID		-869	-864	+
HNF(HNF3, HNF5)		-865	-854	-
GATA1(NF-E1A)		-844	-835	-
GATA1(NF-E1A)		-829	-820	+-
SP1		-803	-794	+
GR		-788	-775	+
HNF(HNF3, HNF5)		-782	-750	+

TABLE 1 (CONTINUED)

Transcription factor		Start	End	Strand
AP1(C-JUN)		-731	-725	+-
PEA3		-731	-726	+-
PU.1(NF-JB)		-730	-725	-
PU.1(NF-JB)		-696	-691	-
AP2		-675	-665	-
AP2		-671	-666	+
H-APF-1		-658	-652	+
NF-kappaB		-658	-647	+
GATA1(NF-E1A)		-631	-626	+-
SP1		-565	-559	-
CAC-bf(htbeta)		-557	-547	+
SP1		-551	-542	-
TFIID		-537	-532	-
NF-E(NF-E1C)		-509	-505	-
CCAT-bf(CBF)		-508	-502	+-
GATA1(NF-E1A)		-508	-504	+
CTF/CBP(NF1)		-507	-502	+
CP2		-506	-502	-
NF-Y		-506	-502	+
CCAAT-bf(CBF)		-464	-458	+-

TABLE 1 (CONTINUED)

Transcription factor		Start	End	Strand
CTF/CBP(NF1)		-464	-459	-
NF-Y		-464	-460	-
SP1		-462	-453	-
SP1		-459	-453	+
AP2		-458	-447	+
PU.1(NF-JB)		-445	-440	+
PEA3		-444	-439	+-
SRY		-439	-428	+
PEA3		-431	-426	+-
CEBP(CEBPA CEBPB)		-420	-412	-
PU.1(NF-JB)		-400	-395	+
PEA3		-399	-394	+-
AP2		-305	-294	-
SP1		-302	-287	+
AP2		-291	-282	-
CEBP(CEBPA CEBPB)		-272	-261	+
PEBP2		-272	-267	+
MYOD		-257	-250	+
E2A		-255	-249	-
SP1		-244	-235	-

TABLE 1 (CONTINUED)

Transcription factor		Start	End	Strand
CAC-bf(htbeta)		-240	-211	+-
SP1		-238	-223	+-
GATA1(NF-E1A)		-225	-219	+
SP1		-225	-215	+
SP1		-221	-209	-
MYB		-211	-204	+
GR		-197	-192	-
SP1		-189	-183	+-
MYB		-174	-167	-
SP1		-166	-152	+-
MYC/MAX		-151	-138	+-
USF 1 and 2 (MYOD) (E-box)		-147	-142	+
AP1(C-JUN)		-131	-121	+
SP1		-100	-86	+-
CEBP(CEBPA CEBPB)		-90	-83	+
GR		-80	-67	-
LXR		-69	-55	+
TFIID(TBP TATA BOX)		-31	-26	+
SIF		-3	2	+

TABLE 1 (CONTINUED)

Transcription factor		Start	End	Strand
SP1		22	31	+

Figure 1 represents a portion of the sequence SEQ ID No. 3. The first nucleotide in the 5' position of the sequence in Figure 1 is also the first nucleotide in the 5' position of one or the nucleic acid sequences SEQ ID No. 1 and SEQ ID No. 3. In the figure, the binding sites for transcription factors are illustrated with boxes which delimit their respective start and end positions, and their respective designations is indicated above each of the corresponding boxes. The numbering of the nucleotides of the sequence represented in Figure 1 was carried out with respect to the transcription start site, numbered "+1", the nucleotide 5' to the nucleotide +1 being itself numbered "-1".

10 The description of the characteristics of the binding sites for each of the transcription factors referred to in Figure 1 and Table 1 can easily be found by persons skilled in the art. A short description of some of them is produced below.

Factor CAC: The characteristics of a binding site for the factor CAC can be found, for example, in the article by Schuele et al., (1988, Nature, Vol. 332: 87-90), entry No. T00077 of the EMBL database, the article by Mantovani et al., (1988, Nucleic Acids Research Vol. 16: 4299-4313), the article by Catala et al., (1989, Nucleic Acid Research, Vol. 17: 3811-3827) and the article by Wang et al., (1993, Mol. Cell Biol., Vol. 13: 5691-5701). The binding of this factor has been shown on the regulatory regions of several genes, including the promoter for the β -globin gene and the gamma-globin gene. This factor appears to act in cooperation with the glucocorticoid receptor.

15
20

Factor C/EBP:**C/EBP- α**

The characteristics of a binding site for the factor C/EBP- α can be found, for example, in the articles corresponding to the following entries in the Medline
5 database: 892020040, 94023981, 96194262, 96003748. This factor inhibits cell proliferation by increasing the level of p21 (WAF-1) due to an increased expression of the gene and to a posttranslational stabilization of p21.

C/EBP- β

The binding characteristics for the factor C/EBP β can be found, for example,
10 in the following entries in the Medline database: 93315489, 91248826, 94193722, 93211931, 92390404, 90258863, 94088523, 90269225 and 96133958. It is a transcription activator which is involved in regulating genes involved in immune and inflammatory responses. It binds specifically to an IL-1 response element in the IL-6 gene. It is believed to play a role in regulating the acute phase of inflammation and
15 in hemopoiesis. The consensus recognition site is as follows: "T(T/G)NNGNAA(T/G)".

Factor c-Myb:

The characteristics of a binding site for the factor c-Myb can be found in the following entries in the Medline database: 91122626, 87092302, 93131991,
20 86261774, 92049347, 90044066, 90265605, 93101590, 94316485 and 90090611. This factor specifically recognizes the sequence "YAAC(G/T)G". It plays a role in controlling the proliferation and differentiation of hematopoietic precursor cells.

Factor CP2:

The characteristics of the binding sites for the factor CP2 can be found, for example, in the articles by Kim et al., (1990), Mol Cell Biol, Vol. 10: 5958-5966 and Lim et al., (1992), J. Biol. Chem. Vol. 268: 18008-18017.

5 **Factor CTF:**

The binding characteristics of the factor CTF can be found, for example, in the following entries in the Medline database: 88319941, 91219459, 86140112, 87237877, 90174951, 89282387, 90151633, 892618136, 86274639, 87064414, 89263791. The factor CTF/NF-I recognizes the following palindromic sequence:
10 "TGGCANNNTGCCA", which is present in viral and cellular promoters and at the origin of replication of type 2 adenoviruses. These proteins are capable of activating transcription and replication. They bind to DNA in the form of a homodimer.

Factor E2A:

The characteristics of a binding site for factor E2A can be found, for example,
15 in the articles corresponding to the following entries in the Medline database: 91160969, 91331308, 91115096, 91117219, 90346284, 89168418, 90150281. This factor binds to a KAPPA-E2 site of the enhancer element of the KAPPA immunoglobulin gene. It forms a heterodimer with the protein ASH1. It belongs to the family of transcription factors of helix-loop-helix type.

20 **Factor GR α :**

The characteristics of a binding site for the GR α factor can be found, for example, in the following entries in the Medline database: 88264449, 93024441, 89091080, 90319784, 92020837, 90381775, 86298392, 91131612, 86092211,

86092206. It is a glucocorticoid receptor which is involved in regulating the expression of eukaryotic genes and which affects the proliferation and differentiation of target tissues. This factor binds to the target site of type "GRE". It is composed of three domains and belongs to the NR3 subfamily of hormone nuclear receptors.

5 **Factor LXR:**

The characteristics of the binding sites for the factor LXR (Liver X Receptor) have been described by Apfel et al., (Moll. Cell. Biol., 1994), Song et al., (Proc. Natl. Acad. Sci USA, 1994) and Willy et al., (Genes Dev., 1995). Oxysterols are the physiological ligand of LXR elements (Janowski et al., Proc. Natl. Acad. Sci USA, 10 1999). Like other receptors, LXR heterodimerizes with RXR (Retinoid X Receptor). Our LXR response elements have been identified at positions -1729/-1714 and -69/-55 (Figure 1).

Factor NF-EIC:

The characteristics of a binding site for the factor NF-EIC can be found, for 15 example, in the following entries in the Medline database: 91266910, 91216113, 91334450, 91203899, 91029498 and 910655813. It is a transcriptional activator which binds to an enhancer element of the α and δ genes of the T cell receptor. It binds to the following consensus sequence: "AGATAG". It belongs to the family of transcription factors of the type GATA.

20 **Factor HNF5:**

Persons skilled in the art may, for example, refer to the article by Grange et al. (1991, Nucleic Acids Res. Vol. 19: 131-139) for this transcription factor.

Factor HNF3B:

Persons skilled in the art may advantageously refer to the article by Overdier et al., (1994, Mol. Cell Biol. Vol. 14: 2755-2766).

Factor Nfkappa-B:

5 Persons skilled in the art may advantageously refer to the articles corresponding to the following entries in the Medline database: 95369245, 91204058, 94280766, 89345587, 93024383, 88248039, 94173892, 91088538, 91239561, 91218850, 92390404, 90156535, 93377072, 92097536, 93309429, 93267517, 92037544, 914266911, 91105848 and 95073993. The factor Nfkappa-B
10 is a heterodimer consisting of a first 50-kDa subunit and a second 65-kDa subunit. Two heterodimers can form a labile tetramer. Its DNA-binding depends on the presence of zinc (Zn^{++}). It can be induced by many agents, such as TNF, PKA or PKC. It is generally a regulator of genes involved in responses to infection, inflammation and stress.

15 **Factor NFY:**

The factor NFY is described, for example, in entry No. P25.208 of the Swissprot database. It is a factor which recognizes a "CCAAT" moiety in promoter sequences such as those of the gene encoding type 1 collagen, of albumin and of β -actin. It is a transcription stimulator.

20 **Factor PEA3:**

Persons skilled in the art may advantageously refer to the articles corresponding to the following entries in the Medline database: 90059931,

90309995, 90291989, 93181246 and 90384794. This transcription factor binds to an "AGGAAG" PEA3 moiety, and can play a regulatory role during embryogenesis.

Factor PEBP2:

Persons skilled in the art may advantageously refer to the articles
5 corresponding to the following entries in the Medline database: 95199266,
94217721, 97188387, 97325753 and 95347606. This factor binds to a
"PYGTGGT" site in many enhancer and promoter elements.

Factor TFIID:

Persons skilled in the art may advantageously refer to the following articles:
10 Fikes et al., (1990, Nature, Vol. 346: 291-294), GILL et al., (1991, Cell, Vol. 65: 333-
340), Hoffmann et al., (1990, Genes Dev. Vol. 4: 1141-1148). This factor plays a
major role in activating the eukaryotic genes which are transcribed by RNA
polymerase II. This factor binds specifically to the TATA promoter element located
close to the transcription start site.

15 **Factor T3R:**

Persons skilled in the art may advantageously refer to the articles
corresponding to the following entries in the Medline database: 92017776,
90242396, 870903752, 91212192. This factor has a strong affinity for
triiodothyroine. It is composed of three domains and belongs to the family of
20 hormone nuclear receptors.

Factor SIF:

Persons skilled in the art may advantageously refer to the article by Wagner
et al., (1990, EMBO J., Vol. 9: 4477-4784). This factor activates the expression of

the c-fos gene.

Factor RAR:

Persons skilled in the art may advantageously refer to the articles corresponding to the following entries in the Medline database: 91216109, 5 92017791, 92127595, 91219411, 92103690, 93321869, 91092269, 91029504, 90242395, 91029504. This factor is a retinic acid receptor. This factor controls cellular functions by directly regulating gene expression. It belongs to the family of hormone nuclear receptors.

Factor PU:

10 Persons skilled in the art may advantageously refer to the articles corresponding to the following entries in the Medline database: 92107189, 93165739, 95317607, 92318913, 92275360, 93028372, 93206099, 90199884, 87257848 and 93275657. This factor binds to the PU box, which is a purine-rich DNA sequence, such as the sequence "GAGGAA", which can act as a lymphoid 15 cell-specific enhancer element.

It is a transcription activation factor which can be specifically involved in the activation or differentiation of macrophages or B cells.

SITE AP1:

The characteristics of a binding site for the transcription factor AP1 can be 20 found in various articles corresponding to the following entries in the Medline database: Numbers 89125693, 89252809, 90318391, 91175677, 911458338, 89313776, 88217909, 911662, 91121514, 89017284, 88070595, 90097934,

88189275, 87301729, 88151062, 90291989, 91330875, 89051877 and 91219459.

This factor binds to an enhancer element of the type "TGA(C/G)TCA.

Factor AP2:

The characteristics of a binding site for the transcription factor AP2 can be
5 found, for example, in the articles corresponding to the following entries in the
Medline database: 90127451, 90174951 and 91009310.

This factor binds to enhancer-type elements, in order to stimulate the
transcription of certain genes. For example, the factor AP2 binds to the following
consensus sequence: "CCCCAGGC".

10 **Factor CCAAT:**

A characteristic of the binding site for the factor CCAAT can be found in the
article by Lum et al., (1990, Mol Cell Biol. Vol. 10: 6709-6717) and in entry No. T
00086 of the EMBL database. This factor has, for example, been shown to be a
stimulator of the transcription of the human hsp70 gene promoter.

15 **Factor GATA-1:**

The characteristics of the binding site for the factor GATA-1 can be found, for
example, in the following entries in the Medline database: 91340773, 91093039,
91266910, 90114418, 89385992, 91268074, 89118131, 91224987, 89218991 and
91081330. It is known to be a "switch" factor in erythroid development. It binds to
20 DNA on the following consensus sequence: "(A/T)GATA(A/G)", in regulatory regions
of globin genes and of other genes expressed in erythroid cells.

Factor MyoD-Myf-3:

The characteristics of a binding site for the factor MyoD-Myf-3 can be found,

for example, in the article by Rosenthal et al., (1990, Nucleic Acids Res., Vol. 18: 6239). This transcription factor induces differentiation of fibroblasts into myoblasts, activates muscle specific promoters and interacts with, and is inhibited by, the protein twist.

5 **Factor MYC/MAX:**

The characteristics of a binding site for the factor Nyc/MAX can be found, for example, in the following entries in the Medline database: 94040733, 93101610, 92229468, 92112037, 93145325, 93026389, 93157390, 92366516, 93145324 and 91173288. This transcription factor binds to DNA nonspecifically, but also binds to
10 DNA specifically by recognizing the sequence CAC[GA]TG. This factor appears to activate the transcription of genes associated with growth.

Factor HNF3:

Persons skilled in the art may advantageously refer to the following entries in the Medline database: 91352065, 91032994, 92345837, 89160814, 91187609,
15 91160974, 91029477, 94301798 and 94218249. This transcription factor acts as an activator of many genes of the liver, such as AFT, albumin and tyrosine aminotransferase genes, and interacts with regulatory regions which are *cis*-acting with respect to these genes.

Factor SRY:

20 Persons skilled in the art may advantageously refer to the articles corresponding to the following entries in the Medline database: 92132550, 95292338, 95112822, 93049201. This factor is responsible, for example, for initiating male sex determination.

Factor Sp1:

Persons skilled in the art may advantageously refer to the articles corresponding to the following entries in the Medline database: 852707437, 89091123, 89039842, 89384647, 85061571, 88111565, 91224491, 91095025, 5 91357479, 91139695. This factor activates the synthesis of messenger RNA from genes containing functional recognition sites; it can interact with G/C base-rich moieties of the promoter of the serotonin receptor gene.

Factor USF:

USF belongs to the helix-loop-helix family of transcription factors that bind the 10 E-box motif (CACGTG) and include inter alia Myc, Mad1, Max, MyoD (Littlewood et al., Oxford University Press, New York, 1pp. 1998). Although it has a well-established role as a transcription activator (Ghosh et al., *Oncogene*, 1997, 14:589-594), it is clear that in some promoter contexts USF does have the capacity to act as a transcription repressor. USF1 has been shown to be important for developmental 15 repression of the LpS1 gene of *Lytechinus pictus* in all cell types except aboral ectoderm cells that do not express USF (Seid et al., *J.Mol Biol*, 1996, 264:7-19). USF1 inhibits autoactivation of the *Xenopus* MyoD gene whose product binds to the E-box (Lun et al., *Cell Growth Differ.*, 1997, 8:275-282) and also represses transcription of the CYP1A1 gene by competing with the stimulatory AhR.Arnt 20 complex for binding to the xenobiotic-responsive element (XRE) that contains an E-box like motif (Takahashi et al., *J Biol Chem*, 1997, 272: 30025-30031). Carter et al. demonstrated that the absence of a strong activation domain in USF1 leads to transcriptional repression of the immunoglobulin heavy-chain (IgH) gene (Carter et

al., *Mol Cell Biol*, 1997, 17:18-23). Truncated and splice variants of USF2 which abrogate E-box activity repress the expression of the major histocompatibility complex class I (Howcroft et al., *Mol Cell Biol* (1999) 19:4788-4797), ATPA (Breen et al., *J Biol Chem*, 1997, 272:10528-10542) and prostaglandin G/H synthase-2
5 genes. Harris et al (*J Biol Chem*, 2000, 275:28539-28548) have provided evidence that the close proximity between the AP-1 and E-box elements on the FGF-BP promoter facilitates transcriptional repression via interactions between USF1, USF2 and the AP-1 binding proteins. In addition, overexpression of human USF decreases AP-1 dependent transcription in murine teratocarcinoma F9 cells (Pognonec et al.,
10 *Oncogene*, 1997, 14:2091-2098) and USF binding to a complex consisting of AP-1 factors, Fra2 and CREB represses the chicken alphaA-crystallin gene (Cvekl et al., 1994, *Mol Cell Biol*, 14:7363-7376). USF1 and USF2 homodimers inhibit transcription of the ribosomal RNA gene (Sirito et al, *Gen Expr*, 1992, 2:231-240). USF binding motifs have also been shown to act as negative regulatory elements in
15 the promoters of the *Protease Nexin-1* (Erno et al, *Mol Cell Neurosci*, 1996, 8:28-37) and the HLA-B (Gobin et al., *J Immunol*, 1999, 163:1428-1434) genes. Interestingly, apolipoprotein CIII, which like ABC1 is involved in lipid metabolism, is also repressed by USF2 (Navantkasattusas et al., *Mol Cell Biol*, 1994, 14: 7331-7339). Finally, USF can both positively and negatively regulate the MLC-2v gene and
20 ribosomal RNA gene transcription. Several studies indicate that USF may also function as a constitutively bound protein that cooperates with basal factors such as TAF_{II}55 (Chiang et al., *Science*, 1995, 267:531-536) or inducible factors such as USA (Meisterernst et al., *Cell*, 1991, 66:981-993), PC5 (Halle et al, *J Biol Chem*,

1995, 270:21307-21311), c-Myc, Max (Harris et al.), CREB and JunD (Cvekl et al.) to mediate transcriptional induction or repression. Thus, USF can function to either activate or repress gene transcription.

Without wishing to be bound by any theory, the applicant thinks that USF may
5 for example inhibit gene transcription by competing for binding to the E-box with transcriptional activators (Lun et al.; Takahashi et al. above cited). However, the applicant has demonstrated in Example 5 that transfection of EBm (mutated E-box) and EBdel (deleted E-box) constructs in RAW cells led to increased rather than decreased ABC1 promoter activity, consistent with E-box mediated gene repression.
10 Furthermore, gel shift assays using antibodies specific to different E-box binding activators failed to demonstrate binding to the E-box motif. Western blot analysis were also performed to determine truncated and splice variants of USF2 that lack the transactivation domain (Liu et al., Sirito et al., et Howcroft et al.) and lead to gene repression. Under these conditions, these USF variants were not detected in
15 RAW cell or 293 cell nuclear extracts.

Because USF can also function as a repressor through specific protein-protein interaction with transcriptional activators that bind DNA motifs distinct from the E-box (Harris et al., Pognonec et al., Cvekl et al., above-cited). As previously described, in the human ABC1 promoter, the E-box motif is flanked by two Sp1 sites
20 and an AP1 motif. It is well known in the art that transcription factors which bind to Sp1 and AP1 have been implicated in the transcriptional regulation of other genes involved in lipid metabolism including apoA-II (Ribeiro et al., J Biol Chem, 265:1216-1225), apoC-III (Ogami et al, J.BiolChem, 1990, 265:9808-9815), chicken

vitellogenin II (Seal et al., Mol Cell Biol, 1991, 11:2704-2717), fatty acid synthase (Casado et al., J Biol Chem, 1999, 274:2009-2013) and the LDL receptor (Sanchez et al., J Biol Chem, 1995, 270:1161-1169) and the LDL receptor related protein (LRP) (Gaeta et al., BBA, 1994, 1219:307-313.)

5 An preferred characteristic of the regulatory nucleic acid according to the invention, and characteristic of the sequence located upstream of the transcription start site, included in both the sequence SEQ ID No. 1 and in the sequence SEQ ID No. 3, is the presence of eight moieties which are characteristic of a putative binding site for the proteins PPAR. The PPARs, also referred to as peroxisome proliferator-
10 activated receptors, which can be of type α , $\delta(\beta)$ and γ , form a subfamily belonging to the family of nuclear receptor genes. All PPARs are activated by fatty acids and derivatives thereof. For example, the PPAR of type α binds to hypolipidemic fibrates, whereas antidiabetic glitazones are ligands for the PPAR of type gamma. Activation of the PPAR of type α induces pleiotropic effects such as the stimulation
15 of lipid oxidation, the impairment of lipoprotein metabolism and the inhibition of vascular inflammation. Activators of PPAR α increase hepatic absorption and esterification of free fatty acids by stimulating the expression of the fatty acid transport protein and of the acyl-CoA synthetase. In the skeletal muscle and the heart, PPAR α increases mitochondrial absorption of free fatty acids, and their
20 oxidation, by stimulating muscle-specific carnitine palmitoyl transferase I. The effect of fibrates on the metabolism of triglyceride-rich lipoproteins is due to the stimulation of lipoprotein lipase, this stimulation being dependent on PPAR α , and to the inhibition of apolipoprotein C-III, whereas the increase in plasmatic cholesterol, in

the form of HDL, depends on an overexpression of apolipoprotein A-I and apolipoprotein A-II.

PPAR α s are also expressed in atherosclerotic adhesions. PPAR α inhibits inducible nitric oxide synthase in macrophages and prevents IL-1-induced
5 cyclooxygenase-2 and IL-6 expression, as well as thrombin-induced endothelin-1 expression, which results from a negative transcriptional regulation of the nuclear factor of the signaling pathways of the nuclear factor NF-KAPPA B and of activating protein-1.

The activation of PPAR α also induces apoptosis in monocyte-derived
10 macrophages, probably by inhibiting the activity of NFKAPPA B. Thus, the pleiotropic effects of the PPAR α activators on the plasmatic lipid profile certainly participate in inhibiting the development of atherosclerosis. PPAR α activators, such as fibrates, inhibit the development of atherosclerosis because of their normolipidemic activities.

15 The presence of eight potential PPAR binding sites (positions of the start and end nucleotides, with respect to the transcription start site, for each of the sites: -1280-1276/-1270-1264, -889-883/-878-872, -584-578/-575-569 and -366-360/-358-352 [also Figure 1]) on a regulatory nucleic acid according to the invention is compatible with the observation according to which the expression of
20 the gene encoding the human ABC1 protein is induced during the differentiation of monocytes into macrophages in vitro. It is also compatible with the prior observation of the regulation of the ABC1 gene by fibrates. It is also compatible with experimental results demonstrating that the expression of the ABC1 gene is

increased when human macrophages are incubated in the presence of acetylated low density lipoproteins (AcLDLs).

Without wishing to be bound by any theory, the applicant thinks that the PPAR binding sites identified according to the invention on the regulatory nucleic acid of sequence SEQ ID No. 1 are highly involved in the tissue specific regulation, and in the specific regulation of the cellular metabolic situation, of the ABC1 gene, and that as a result, a regulatory sequence which comprises at least 4, for example, at least 5, 6, 7 or all of the 8 PPAR binding sites (Figure 1)(of the sequence SEQ ID No. 1, and which also comprises a basic promoter element, is useful as a regulatory sequence for a polynucleotide whose expression is desired in the liver, the lungs, the adrenal glands, the monocytes/macrophages, the placenta or the fetal tissues, or for a polynucleotide whose expression is desired in response to a specific stimulation of the cell, in relation with cholesterol metabolism, such as the presence, in the cellular environment, of acetylated low density lipoproteins (Ac LDLs).

In addition, it has been shown according to the invention that a regulatory nucleic acid for the human ABC1 gene, as defined above, which comprises all the abovementioned PPAR sites, is capable of regulating the expression of a coding sequence placed under its control, in a manner which is dependent on the presence of cholesterol in the cellular environment. The results are presented in Example 4 below.

As previously mentioned, the invention concerns a nucleic acid comprising a polynucleotide which has at least 20 consecutive nucleotides of either of the

nucleotide sequences SEQ ID No. 1 and SEQ ID No. 2, as well as a nucleic acid of complementary sequence.

The nucleic acids comprising one or more "biologically active" fragments of either of the sequences SEQ ID No. 1 and SEQ ID No. 2 are encompassed in the definition above. Persons skilled in the art can easily obtain biologically active fragments of these sequences by referring, for example, to Table 1 below and to Figure 1, in which are presented the various moieties which are characteristic of the regulatory sequence for the ABC1 gene. Persons skilled in the art may thus obtain such biologically active fragments by totally or partially chemically synthesizing the corresponding polynucleotides or by using restriction endonucleases to obtain desired DNA fragments, the restriction sites present on the sequences SEQ ID No. 1 and SEQ ID No. 1 being able to be easily found from the sequence information, with the aid of current restriction mapping software such as GCG version 9.1 module map.

The production of nucleic acid fragments determined using restriction endonucleases is described, for example, in the work by Sambrook et al. (1989).

The invention thus also relates to a nucleic acid, as defined above, which is capable of modulating the transcription of a polynucleotide placed under its control.

According to one embodiment, a biologically active fragment of a transcription-regulating acid according to the invention comprises the basic promoter (TATA box and homeobox) ranging from the nucleotide at position -1 to the nucleotide at position -300, with respect to the transcription start site, the first

nucleotide transcribed being the nucleotide in position 2894 of the nucleotide sequence SEQ ID No. 1.

According to a second embodiment, a biologically active fragment of a transcription-regulating nucleic acid according to the invention comprises both the
5 basic promoter and the proximal regulatory elements, ranging from the nucleotide at position -1 to the nucleotide at position -600, with respect to the transcription start site, the first nucleotide transcribed being the nucleotide in position 2894 of the nucleotide sequence SEQ ID No. 1.

According to a third embodiment, a biologically active fragment of a
10 transcription-regulating nucleic acid according to the invention comprises besides the basic promoter (core promoter) and the proximal 200 bp of the ABC1 gene promoter which is rich in binding sites for transcription factors, *i.e.*, Sp1, AP1, LXR, and E-box, that are likely involved in modulating human ABC1 gene expression.

According to a fourth embodiment, such a biologically active fragment of a
15 transcription-regulating acid according to the invention also comprises, besides the basic promoter (core promoter) and the proximal regulatory elements, other regulatory elements such as the various PPAR α sites, and stretches from the nucleotide at position -1 to the nucleotide at position -2894, with respect to the transcription start site, the first nucleotide transcribed being the nucleotide in
20 position 2894 of the nucleotide sequence SEQ ID No. 1.

According to a fifth embodiment, such a biologically active fragment of a transcription-regulating acid according to the invention, which also comprises, besides the basic promoter (core promoter) and the proximal regulatory elements,

other regulatory elements such as the various PPAR α sites, stretches from the nucleotide at position +120 to the nucleotide at position -995, with respect to the transcription start site, the first nucleotide transcribed being the nucleotide in position 2894 of the nucleotide sequence SEQ ID No. 1.

- 5 According to a sixth embodiment, a biologically active fragment of a transcription-regulating nucleic acid according to the invention comprises a region ranging from the nucleotide at position +108 to the nucleotide at position -2228, with respect to the transcription start site as set forth in sequence SEQ ID No. 1.

10 **EXONS 1A and 1B AND INTRONS 1A and 1B**

The applicant has also identified the nucleotide sequences located downstream of the transcription start site, and corresponding respectively, from the 5' end to the 3' end, to exon 1A, intron 1A, exon 1B and intron 1B of the human gene encoding the ABC1 protein.

- 15 More precisely, exon 1A, which is 221 nucleotides long, starts at the nucleotide at position 2894 of the sequence SEQ ID No. 1 and ends at the nucleotide at position 3114 of the sequence SEQ ID No. 1. Exon 1A is identified as the sequence SEQ ID No. 4.

- Exon 1B, which is 109 nucleotides long, starts at the nucleotide at position
20 100 and ends at the nucleotide at position 258 of the sequence SEQ ID No. 2. Exon 1B is identified as the sequence SEQ ID No. 5.

Intron 1A has been partially sequenced. The 5' end of intron 1A starts at the nucleotide at position 3115 and ends at the nucleotide at position 3231 of the

nucleotide sequence SEQ ID No. 1, and is also defined as the sequence SEQ ID No. 6. The 3' end of intron 1A starts at the nucleotide at position 1 and ends at the nucleotide at position 99 of the nucleotide sequence SEQ ID No. 2, and is also identified as the sequence SEQ ID No. 7.

5 Intron 1B has been partially sequenced. The 5' end of intron 1B starts at the nucleotide at position 259 and ends at the nucleotide at position 357 of the sequence SEQ ID No. 2. This sequence is also identified as the sequence SEQ ID No. 8.

 Exon 1B contains the start of the open reading frame of the human ABC1
10 gene, the nucleotide A of the ATG codon being located in start position at position 94 of the sequence SEQ ID No. 5. Exon 1B encodes the polypeptide of sequence SEQ ID No. 9.

 Exons 1A and 1B, as well as introns 1A and 1B, can contain elements for regulating the expression of the ABC1 gene, for example, elements of enhancer
15 type and/or elements of silencer type.

 Consequently, a transcription-regulating nucleic acid according to the invention can also contain, besides biologically active fragments of the sequence SEQ ID No. 1, nucleotide fragments, or even all, of the sequences SEQ ID No. 2 to SEQ ID No. 8.

20 The nucleotide sequences SEQ ID No. 1 to SEQ ID No. 8, as well as fragments thereof, can, for example, be used as nucleotide probes or primers for detecting the presence of at least one copy of the ABC1 gene in a sample, or for amplifying a given target sequence in the regulatory sequence for the ABC1 gene.

An individual of the invention is thus a nucleic acid having at least 80% nucleotide identity with a nucleic acid as defined above, for example, originating from one of the sequences SEQ ID No. 1 to SEQ ID No. 8.

The invention also concerns a nucleic acid which hybridizes, under high stringency conditions, with any one of the nucleic acids according to the invention, for example, a nucleic acid originating from a sequence chosen from the sequences SEQ ID No. 1 to SEQ ID No. 8.

The invention also relates to a nucleic acid as defined above and also further characterized in that it is capable of modifying the transcription of a polynucleotide of interest placed under its control.

According to a first aspect, such a nucleic acid is capable of activating the transcription of the polynucleotide of interest placed under its control.

According to a second aspect, a regulatory nucleic acid according to the invention can be characterized in that it is capable of inhibiting the transcription of the polynucleotide of interest placed under its control.

For example, a transcription-regulating nucleic acid according to the invention, when it is suitably located with respect to a polynucleotide of interest whose expression is desired, will allow the transcription of said polynucleotide of interest, either constitutively or inducibly.

The inducible nature of the transcription initiated by a regulatory nucleic acid according to the invention can be conferred by one or more of the regulatory elements that it contains, for example, the presence of one or more PPAR α sites,

such as at least 4 PPAR α sites, or at least 5, 6, 7 PPAR α sites or the 8 PPAR α sites of the sequence SEQ ID No. 1 or SEQ ID No. 3.

In addition, a tissue specific expression of the polynucleotide of interest can be sought by placing this polynucleotide of interest under the control of a regulatory
5 nucleic acid according to the invention which is capable, for example, of initiating the transcription of this polynucleotide of interest specifically in certain categories of cells, for example, cells from the liver, placenta cells or macrophages.

In general, a regulatory nucleic acid according to the invention can comprise one or more "discrete" regulatory elements, such as enhancer and silencer
10 elements. For example, such a regulatory nucleic acid can comprise one or more potential transcription factor binding sites as defined in Figure 1.

A regulatory acid according to the invention also encompasses a sequence which does not comprise the basic promoter, i.e. the sequence ranging from the nucleotide at position -1 to the nucleotide at position -300, with respect to the
15 transcription start site.

Such a regulatory nucleic acid will then generally comprise a so-called "heterologous" basic promoter, i.e. a polynucleotide comprising a "TATA" box and a "homeobox", which does not originate from the regulatory nucleic acid for the ABC1 gene.

20 A transcription-regulating nucleic acid comprising all or part of the sequence SEQ ID No. 1 which has been modified, for example, by addition, deletion or substitution of one or more nucleotides, also forms part of the invention. Such

modifications can modify the transcriptional activity by causing an increase, or on the contrary a decrease, in the activity of the promoter or of the regulatory element.

Such a modification can also affect the tissue specificity of the promoter or of the regulatory element. Thus, for example, a regulatory nucleic acid according to
5 the invention can be modified in order to stimulate transcription in only one of the tissues in which it is naturally expressed.

A transcription-regulating acid according to the invention can also be modified and be made inducible by a specific compound, for example, by creating, in the sequence, a site which is inducible by a given therapeutic compound.

10 The modifications in a sequence which comprises all or part of the sequence SEQ ID No. 1 and which comprises the promoter or a regulatory element can be carried out using methods that are well known to persons skilled in the art, such as mutagenesis. The activity of the modified promoter or regulatory element can then be tested, for example, by cloning the modified promoter upstream of a reporter
15 gene, by transfecting the resulting DNA construct into a host cell and by measuring the level of expression of the reporter gene in the transfected host cell. The activity of the modified promoter can also be analyzed *in vivo* in transgenic animals. It is also possible to construct libraries of modified fragments, which can be screened using functional tests in which, for example, only the promoters or the regulatory
20 elements having the desired activity will be selected.

Such assays can be based, for example, on the use of reporter genes which confer resistance to given compounds, for example, to antibiotics. The selection of cells which have a regulatory nucleic acid/reporter gene construct, and which

contain a promoter or a regulatory element with the desired modification, can then be isolated by culturing the transformed host cells having such a construct, in the presence of the given compound, for example, of the given antibiotic.

The reporter gene can also encode any easily detectable protein, for example, an optically detectable protein such as luciferase.

Consequently, an individual of the invention is also a nucleic acid comprising:

a) a transcription-regulating nucleic acid as defined above; and

b) a polynucleotide of interest encoding a polypeptide or a nucleic acid of interest.

According to a first aspect, the polynucleotide of interest whose transcription is desired encodes a protein or a peptide. The protein can be of any nature, for example, a protein of therapeutic interest, including cytokines, structural proteins, receptors or transcription factors. For example, when transcription specifically in certain tissues is desired, such as for example, in liver, macrophage or placenta cells, the transcription-regulating nucleic acid will advantageously comprise a nucleic acid ranging from the nucleotide at position -1 to the nucleotide at position -1318, with respect to the transcription start site of the sequence SEQ ID No. 1 or SEQ ID No. 3.

In this case, the polynucleotide of interest will encode a gene involved in combating inflammation, such as a cytokine receptor, or a superoxide dismutase. If an antitumoral effect is desired, then stimulation of the number and the activation of cytotoxic T lymphocytes specific for a given tumoral antigen will be sought.

Alternatively, the regulatory nucleic acid advantageously comprises a nucleic acid ranging from the nucleotide at position +108 to the nucleotide -2228 , with respect to the transcription start site of the sequence SEQ ID NO: 1.

In another embodiment, a regulatory nucleic acid according to the invention
5 will be used in combination with a polynucleotide of interest encoding the ABC1 protein.

As already mentioned, the polynucleotide of interest can also encode a nucleic acid, such as an antisense nucleic acid specific for a gene the inhibition of whose translation is desired.

10 According to another aspect, the polynucleotide of interest, whose transcription is regulated by the regulatory nucleic acid, is a reporter gene, such as any gene encoding a detectable protein.

Among the exemplified reporter genes, mention may be made for example, of the luciferase, the β -galactosidase (LacZ) or the chloramphenicol acetyl transferase
15 (CAT) gene, or any gene encoding a protein which confers resistance to a specific compound, such as to an antibiotic.

RECOMBINANT VECTORS

For the purposes of the present invention, "vector" will be intended to mean a
20 linear or circular RNA or DNA molecule which is indifferently in the single-stranded or double-stranded form.

According to a first embodiment, a recombinant vector according to the invention is used to amplify the regulatory nucleic acid according to the invention, which is inserted therein after transformation or transfection of the desired host cell.

According to a second embodiment, they are expression vectors comprising,
5 besides a regulatory nucleic acid in accordance with the invention, sequences whose expression is sought in a host cell or in a given multicellular organism.

According to an advantageous embodiment, a recombinant vector according to the invention will, for example, comprise the following elements:

- (1) a regulatory nucleic acid according to the invention;
- 10 (2) a polynucleotide of interest comprising a coding sequence included in the nucleic acid to be inserted into such a vector, said coding sequence being placed in frame with the regulatory signals described in (1); and
- (3) suitable transcription start and stop sequences.

In addition, the recombinant vectors according to the invention may include
15 one or more origins of replication in the host cells in which their amplification or their expression is desired, markers or selection markers.

By way of examples, the bacterial promoters may be the promoters LacI or LacZ, the T3 or T7 bacteriophage RNA polymerase promoters, or the lambda phage PR or PL promoters.

20 The promoters for eukaryotic cells may comprise the thymidine kinase promoter of the virus HSV or the mouse metallothioneine-L promoter.

In general, for the choice of a suitable promoter, persons skilled in the art may advantageously refer to the abovementioned work by Sambrook et al. (1989) or to the techniques described by Fuller et al. (1996).

When the expression of the genomic sequence of the ABC1 gene is desired,
5 vectors which are capable of including large insertion sequences will, for example, be used. In this specific embodiment, bacteriophage vectors, such as the P1 bacteriophage vectors like the vector p158 or the vector p158/neo8 described by Sternberg (1992, 1994), will, for example, be used.

Bacterial vectors according to the invention can be, for example, the vectors
10 pBR322(ATCC37017) or vectors such as pAA2233 (Pharmacia, Uppsala, Sweden) and pGEM1 (Promega Biotech, Madison, WI, USA).

Mention may also be made of other commercialized vectors, such as the vectors pQE70, pQE60, pQE9 (Qiagen), psiX174, pBluescript SA, pNH8A, pNH16A, pNH18A, pNH46A, pWLNEO, pSV2CAT, pOG44, pXTI, pSG(Stratagene).

15 It can also be the recombinant vector PXP1 described by Nordeen SK et al. (1988, BioTechniques, 6: 454-457).

They can also be vectors of *Baculovirus* type, such as the vector pVL1392/1393 (Pharmingen) used for transfecting cells of the Sf9 line (ATCC No. CRL 1711) derived from *Spodoptera frugiperda*.

20 They can also be adenoviral vectors, such as the type 2 or 5 human adenovirus.

A recombinant vector according to the invention can also be a retroviral vector or an adeno-associated vector (AAV). Such adeno-associated vectors are

described, for example, by Flotte et al. (1992), Samulski et al. (1989) or McLaughlin BA et al. (1996).

To enable the expression of a polynucleotide of interest under the control of a regulatory nucleic acid according to the invention, the polynucleotide construct
5 comprising the regulatory sequence and the coding sequence may be introduced into a host cell. The introduction of such a polynucleotide construct according to the invention into a host cell can be carried out *in vitro*, according to the techniques well known to persons skilled in the art for transforming or transfecting cells, either in primary culture or in the form of cell lines. The introduction of the polynucleotides
10 according to the invention can also be carried out *in vivo* or *ex vivo*, for preventing or treating diseases linked to a deficiency in the reverse transport of cholesterol.

To introduce the polynucleotides or the vectors into a host cell, persons skilled in the art may advantageously refer to various techniques, such as the calcium phosphate precipitation technique (Graham et al., 1973, Chen et al., 1987),
15 DEAE Dextran (Gopal, 1985), electroporation (Tur-Kaspa, 1996; Potter et al., 1984), direct microinjection (Harland et al., 1985) or DNA-loaded liposomes (Nicolau et al., 1982, Fraley et al., 1979).

Once the polynucleotide has been introduced into the host cell, it can be stably integrated into the genome of the cell. The integration can be carried out at a
20 specific place in the genome, by homologous recombination, or it can be randomly integrated. In some embodiments, the polynucleotide can be stably maintained in the host cell in the form of an episome fragment, the episome comprising

sequences which allow the latter to be maintained and replicated, either independently or in synchrony with the cell cycle.

According to one embodiment, one method for introducing a polynucleotide according to the invention into a host cell, such as a host cell originating from a mammal, *in vivo*, comprises a preparation comprising a pharmaceutically compatible vector and a "naked" polynucleotide according to the invention, placed under the control of suitable regulatory sequences, are introduced by local injection into the chosen tissue for example, a smooth muscle tissue, the "naked" polynucleotide being absorbed by the cells of this tissue.

Compositions for use *in vitro* and *in vivo* comprising "naked" polynucleotides are described for example, in PCT Application No. WO 95/11307 (Pasteur Institute, Inserm, University of Ottawa) and in the articles by Tacson et al. (1996) and by Huygen et al. (1996).

According to one specific embodiment of the invention, a composition is provided for producing a protein of interest *in vivo*. This composition comprises a polynucleotide encoding the polypeptide of interest, placed under the control of a regulatory sequence according to the invention, in solution in a physiologically acceptable vector.

The amount of vector which is injected into the chosen host organism varies according to the site of injection. By way of indication, between approximately 0.1 and approximately 100 µg of the regulatory sequence/coding sequence polynucleotide construct can be injected into the body of an animal.

When the regulatory nucleic acid according to the invention is located, on the polynucleotide construct (or vector), so as to control the transcription of a sequence comprising an open reading frame encoding the ABC1 protein, the vector may be injected into the body of a patient likely to develop a disease linked to a deficiency in
5 the reverse transport of cholesterol, or who has already developed this disease, for example, a patient with a predisposition for Tangier disease, or who has already developed the disease.

Consequently, the invention also concerns a pharmaceutical composition intended for preventing, or for treating individuals affected by, a dysfunction of the
10 reverse transport of cholesterol, comprising a regulatory nucleic acid according to the invention and a polynucleotide of interest encoding the ABC1 protein, in combination with one or more physiologically compatible excipients.

Advantageously, such a composition will comprise the regulatory nucleic acid defined by either of the sequences SEQ ID No. 1 and SEQ ID No. 2, or a biologically
15 active fragment of this regulatory nucleic acid.

A subject of the invention is also a pharmaceutical composition intended for preventing, or for treating individuals affected by, a dysfunction of the reverse transport of cholesterol, comprising a recombinant vector as defined above in association with one or more physiologically compatible excipients.

20 The invention also relates to the use of a polynucleotide construct in accordance with the invention which comprises a regulatory nucleic acid for the ABC1 gene, as well as a sequence encoding the ABC1 protein, for manufacturing a medicinal product intended for preventing atherosclerosis in various forms or, for

example, for treating individuals affected by a dysfunction of the reverse transport of cholesterol.

The invention also relates to the use of a recombinant vector according to the invention, comprising, besides a regulatory nucleic acid of the invention, a nucleic acid encoding the ABC1 protein, for manufacturing a medicinal product intended for preventing atherosclerosis in various forms or, for example, for treating individuals affected by a dysfunction of the reverse transport of cholesterol.

Vectors which are useful in somatic gene therapy methods and compositions containing such vectors

The present invention also concerns a novel therapeutic approach for treating pathologies linked to cholesterol transport. It proposes an advantageous solution to the drawbacks of the prior art, by demonstrating the possibility of treating pathologies, for example, pathologies linked to cholesterol transport, by gene therapy, by transferring and expressing, *in vivo*, a polynucleotide construct comprising, besides a regulatory nucleic acid according to the invention, a sequence encoding an ABC1 protein involved in cholesterol transport and metabolism. The invention also offers a simple means which allows specific and effective treatment of the associated pathologies such as for example, atherosclerosis.

Gene therapy consists in correcting a deficiency or an abnormality (mutation, aberrant expression, etc.) or in effecting the expression of a therapeutic protein of interest by introducing genetic information into the affected cell or organ. This genetic information can be introduced either *ex vivo*, into a cell extracted from the

organ, the modified cell then being reintroduced into the body, or directly in vivo into the appropriate tissue. In this second case, various techniques exist, among which diverse techniques of transfection involving complexes of DNA and of DEAE-dextran (Pagano et al., J. Virol. 1 (1967) 891), of DNA and of nuclear proteins (Kaneda et al., Science 243 (1989) 375), and of DNA and of lipids (Felgner et al., PNAS 84 (1987) 7413), the use of liposomes (Fraley et al., J. Biol. Chem. 255 (1980) 10431), etc. More recently, the use of viruses as vectors for transferring genes has appeared as a promising alternative to these physical techniques of transfection. In this respect, various viruses have been tested for their capacity to infect certain cell populations. For example, retroviruses (RSV, HMS, MMS, etc.), the HSV virus, adeno-associated viruses and adenoviruses.

The present invention thus also relates to a novel therapeutic approach for treating pathologies linked to cholesterol transport, consisting in transferring and in expressing, in vivo, genes encoding ABC1 placed under the control of a regulatory acid according to the invention. The applicant has now advantageously shown that it is possible to construct recombinant viruses which contain a DNA sequence comprising a regulatory nucleic acid according to the invention and a sequence encoding an ABC1 protein involved in cholesterol metabolism, and to administer these recombinant viruses in vivo, and that this administration enables an expression of a biologically active ABC1 protein in vivo, which is stable and effective, and which is without cytopathological effect.

Adenoviruses constitute vectors which are efficient for transferring and expressing the ABC1 gene. For example, the use of recombinant adenoviruses as

vectors makes it possible to obtain levels of expression of the gene of interest which are sufficiently high to produce the desired therapeutic effect. Other viral vectors, such as retroviruses or adeno-associated viruses (AAVs), which enable a stable expression of the gene, are also claimed.

5 The present invention thus offers a novel approach for treating and preventing cardiovascular and neurological pathologies linked to the abnormalities in cholesterol transport.

A subject of the invention is thus a defective recombinant virus comprising a regulatory nucleic acid according to the invention and a nucleic acid sequence
10 encoding an ABC1 protein involved in cholesterol metabolism.

The invention also relates to the use of such a defective recombinant virus for preparing a pharmaceutical composition intended for treating and/or for preventing cardiovascular diseases.

The present invention also concerns the use of cells which are genetically
15 modified, ex vivo with a virus as described above, or of producer cells for such viruses, which are implanted in the body, enabling a protracted and effective in vivo expression of a biologically active ABC1 protein.

The present invention shows that it is possible to incorporate a DNA sequence encoding ABC1, under the control of a regulatory nucleic acid as defined
20 above, into a viral vector, and that these vectors make it possible to effectively express a biologically active mature form. For example, the invention shows that the in vivo expression of ABC1 can be obtained by direct administration of an

adenovirus or by implantation of a producer cell or a cell which is genetically modified with an adenovirus or with a retrovirus incorporating such a DNA.

The present invention is also advantageous because it makes it possible to induce an expression of ABC1 which is controlled and without harmful effects, in
5 organs which the expression of this protein does not normally concern. For example, a significant release of the ABC1 protein is obtained by implantation of cells which produce vectors of the invention, or which are infected ex vivo with vectors of the invention.

The cholesterol transporter activity produced in the context of the present
10 invention can be of human or animal ABC1 type. The nucleic acid sequence used in the context of the present invention can be a cDNA, a genomic DNA (gDNA) an RNA (in the case of retroviruses) or a hybrid construct consisting for example, of a cDNA into which one or more introns would be inserted. It can also be synthetic or semisynthetic sequences. For example, a cDNA or a gDNA is used. For example,
15 the use of a gDNA allows better expression in human cells. To allow their incorporation in a viral vector according to the invention, these sequences are advantageously modified, for example, by site-directed mutagenesis, for example, for inserting suitable restriction sites. The sequences described in the prior art are not in fact constructed for a use according to the invention, and prior adjustments
20 may prove to be necessary, so as to obtain substantial expressions. In the context of the present invention, use of a nucleic acid sequence encoding a human ABC1 protein is one example. Moreover, it is also possible to use a construct encoding a derivative of these ABC1 proteins. A derivative of these ABC1 proteins comprises,

for example, any sequence which is obtained by mutation, deletion and/or addition, with respect to the native sequence, and which encodes a product which conserves cholesterol transporter activity. These modifications can be carried out by the techniques known to persons skilled in the art (see the general molecular biology techniques below). The biological activity of the derivatives thus obtained can then be easily determined, as indicated, for example, in the examples describing the measurement of the afflux of cholesterol from cells. For the purposes of the invention, the derivatives can also be obtained by hybridization using nucleic acid libraries, using the native sequence or a fragment of this sequence as probe.

10 These derivatives are, for example, molecules with greater affinity for their binding sites, molecules having greater resistance to protease, or molecules with a greater therapeutic efficacy or fewer side effects, or optionally, novel biological properties. The derivatives also include the modified DNA sequences which allow improved expression in vivo.

15 In a first embodiment, the present invention concerns a defective recombinant virus comprising a regulatory nucleic acid according to the invention and a cDNA sequence encoding an ABC1 protein involved in cholesterol transport and metabolism. In another embodiment of the invention, the DNA sequence is a gDNA sequence. The cDNA sequence which encodes the ABC1 protein, and which
20 can be used in a vector according to the invention, is advantageously the sequence SEQ ID No. 10.

The vectors of the invention can be prepared from various types of virus. For example, vectors derived from adenoviruses, adeno-associated viruses (AAV),

herpesviruses (HSV) or retroviruses are used. It is also advantageous to use an adenovirus for a direct administration or for modifying, ex vivo, cells intended for implanting, or a retrovirus for implanting producer cells.

The viruses according to the invention are defective, i.e. they are incapable of
5 replicating autonomously in the target cell. Generally, the genome of the defective viruses used in the context of the present invention is thus lacking at least the sequences required for replication of said virus in the infected cell. These regions can be either removed (totally or partially), or rendered nonfunctional or substituted with other sequences, and, for example, with the nucleic acid sequence encoding
10 the ABC1 protein. For example, however, the defective virus conserves the sequences in its genome which are required for encapsidation of viral particles.

In regards to adenoviruses, various serotypes, whose structure and properties vary somewhat, have been characterized. For example, among these serotypes, in the context of the present invention, type 2 or 5 human adenoviruses
15 (Ad 2 or Ad 5) or adenoviruses of animal origin (see application WO 94/26914) may be used. Among the adenoviruses of animal origin which can be used in the context of the present invention, mention may be made of the adenoviruses of canine, bovine, murine (example: Mav1, Beard et al., Virology 75 (1990) 81), ovine, porcine, avian or simian (example: SAV) origin. For example, the adenovirus of animal origin
20 may be a canine adenovirus, such as a CAV2 adenovirus [Manhattan or A26/61 strain (ATCC VR-800) for example]. For example, in the context of the invention, adenoviruses of human or canine or mixed origin may be used. In addition, the defective adenoviruses of the invention may also comprise ITRs, a sequence

enabling the encapsidation and the sequence encoding the ABC1 protein placed under the control of a nucleic acid according to the invention. In one embodiment of the present invention, in the genome of the adenoviruses of the invention, the E1 region at least is rendered nonfunctional. In another embodiment, in the genome of the adenoviruses of the invention, the E1 gene and at least one of the genes E2, E4 and L1-L5 are nonfunctional. The viral gene under consideration can be rendered nonfunctional by any technique known to persons skilled in the art, and, for example, by total elimination, substitution, partial deletion, or addition of one or more bases in the gene(s) under consideration. Such modifications can be obtained in vitro (on isolated DNA) or in situ, for example, by means of the techniques of genetic engineering or by treatment with mutagenic agents. Other regions can also be modified, and, for example, the E3 region (WO 95/02697), E2 (WO 94/28938), E4 (WO 94/28152, WO 94/12649, WO 95/02697) and L5 (WO 95/02697). According to one embodiment, the adenovirus according to the invention comprises a deletion in the E1 and E4 regions, and the sequence encoding ABC1 is inserted into the inactivated E1 region. According to another embodiment, it comprises a deletion in the E1 region, into which are inserted the E4 region and the sequence encoding ABC1 (French Patent Application FR 94 13355).

The defective recombinant adenoviruses according to the invention can be prepared by any technique known to persons skilled in the art (Levrero et al., Gene 101 (1991) 195, EP 185 573; Graham, EMBO J. 3 (1984) 2917). For example, they can be prepared by homologous recombination between an adenovirus and a plasmid carrying, inter alia, the DNA sequence encoding the ABC1 protein. The

homologous recombination takes place, for example, after cotransfection of said adenovirus and plasmid into a suitable cell line. The cell line used should, for example, (i) be transformable by said elements, and (ii) have the sequences which are capable of complementing the portion of the genome of the defective
5 adenovirus, such as in an integrated form to avoid the risks of recombination. By way of example of a line, mention may be made of the 293 human embryonic kidney line (Graham et al., J. Gen. Virol. 36 (1977) 59), which comprises, for example, integrated into its genome, the left-hand portion of the genome of an adenovirus Ad5 (12%) or lines which are capable of complementing the E1 and E4 functions, as
10 described, for example, in applications No. WO 94/26914 and WO 95/02697.

Next, the adenoviruses which have multiplied may be recovered and purified according to the conventional techniques of molecular biology, as illustrated in the examples.

As regards the adeno-associated viruses (AAVs), they are relatively small
15 DNA viruses, which integrate into the genome of the cells that they infect, in a stable and site-specific way. They are capable of infecting a broad spectrum of cells, without inducing effects on cell growth, morphology or differentiation. Moreover, they do not seem to be involved in pathologies in humans. The AAV genome has been cloned, sequenced and characterized. It comprises approximately 4700 bases
20 and contains, at each end, an inverted repeat region (ITR) of 145 bases approximately which serves as an origin of replication for the virus. The rest of the genome is divided into two essential regions carrying the encapsidation functions: the left-hand portion of the genome, which contains the rep gene which is involved

in viral replication and expression of viral genes; the right-hand portion of the genome, which contains the cap gene which encodes the capsid proteins of the virus.

The use of AAV-derived vectors for transferring genes in vitro and in vivo has
5 been described in the literature (see, for example, WO 91/18088; WO 93/09239;
US 4,797,368, US 5,139,941, EP 488 528). These documents describe various
AAV-derived constructs, in which the rep and/or cap genes are deleted and replaced
with a gene of interest, and their use for transferring said gene of interest in vitro (on
culture cells) or in vivo (directly in an organism). However, none of these documents
10 describes or suggests using a recombinant AAV for transferring and expressing an
ABC1 protein in vivo or ex vivo, or the advantages of such a transfer. The defective
recombinant AAVs according to the invention can be prepared by cotransfecting,
into a cell line which has been infected with a human helper virus (for example, an
adenovirus), a plasmid containing the sequence encoding the ABC1 protein,
15 bordered by two AAV inverted repeat regions (ITR), and a plasmid carrying the AAV
encapsidation genes (rep and cap genes). The recombinant AAVs produced may
then be purified by conventional techniques.

As regards herpes viruses and retroviruses, the construction of recombinant
vectors has been widely described in the literature: see, for example, Breakfield et
20 al., New Biologist 3 (1991) 203; EP 453242, EP 178220, Bernstein et al. Genet.
Eng. 7 (1985) 235; McCormick, BioTechnology 3 (1985) 689, etc.

For example, retroviruses are integrating viruses which infect dividing cells.
The retrovirus genome comprises two LTRs, an encapsidation sequence and three

coding regions (gag, pol and env). In retrovirus-derived recombinant vectors, the gag, pol and env genes are generally totally or partially deleted and replaced with a heterologous nucleic acid sequence of interest. These vectors can be prepared from various types of retrovirus, such as for example, MoMuLV ("murine moloney leukemia virus"; also referred to as MoMLV), MSV ("murine moloney sarcoma virus"), HaSV ("harvey sarcoma virus"), SNV ("spleen necrosis virus"), RSV ("rous sarcoma virus") or the Friend virus.

To construct recombinant retroviruses comprising a sequence encoding the ABC1 protein, placed under the control of a regulatory nucleic acid according to the invention, a plasmid comprising, for example, the LTRs, the encapsidation sequence and said coding sequence is generally constructed, and then used for transfecting a so-called encapsidation cell line which is capable of providing, in trans, the retroviral functions deficient in the plasmid. Generally, encapsidation lines are thus capable of expressing gag, pol and env genes. Such encapsidation lines have been described in the prior art, for example, the line PA317 (US 4,861,719); the line PsiCRIP (WO 90/02806) and the line GP+envAm-12 (WO 89/07150). Moreover, the recombinant retroviruses can comprise modifications in the LTRs so as to eliminate transcriptional activity, as well as extended encapsidation sequences comprising a portion of the gag gene (Bender et al., J. Virol. 61 (1987) 1639). The recombinant retroviruses produced may then be purified by conventional techniques.

To implement the present invention, it is advantageous to use a defective recombinant adenovirus. The results given below indeed demonstrate the

properties of adenoviruses which are advantageous for expressing, in vivo, a protein having cholesterol transport activity. The adenoviral vectors according to the invention are advantageous for direct administration in vivo of a purified substance or for transforming cells, for example, autologous cells, ex vivo, with a view to
5 implanting them. In addition, the adenoviral vectors according to the invention also have advantages, such as, for example, their very high infection efficiency, which makes it possible to carry out infections using small volumes of viral suspension.

According to another embodiment of the invention, a producer line for retroviral vectors containing a regulatory nucleic acid according to the invention and
10 the sequence encoding the ABC1 protein is used for an in vivo implantation. The lines which can be used to this end are, for example, the cells PA317 (US 4,861,719), PsiCrip (WO 90/02806) and GP+envAm-12 (US 5,278,056), which are modified to allow production of a retrovirus containing a nucleic acid sequence encoding an ABC1 protein according to the invention. For example, totipotent stem
15 cells, which are precursors of the blood cell lines, can be sampled and isolated from the individual. These cells, which are put into culture, can then be transfected with the retroviral vector containing the sequence encoding the ABC1 protein, under the control of its own promoter. These cells are then reintroduced into the individual. The differentiation of these cells will be the origin of blood cells which express the
20 ABC1 protein, for example, the origin of monocytes which, when transformed into macrophages, participate in removing cholesterol from the arterial wall. These macrophages expressing the ABC1 protein will have an increased capacity for

metabolizing excess cholesterol, and will make it available at the cell surface for its removal by the primary acceptors of membrane cholesterol.

In one embodiment of the present invention, in the vectors of the invention, the sequence encoding the ABC1 protein is placed under the control of a regulatory
5 nucleic acid according to the invention, comprising the regulatory elements which allow its expression in the infected cells and, for example, the regulatory elements of type PPAR.

In another embodiment, the vectors of the invention comprise the sequence encoding the ABC1 protein which is placed under the control of a regulatory nucleic
10 acid comprising a region ranging from the nucleotide at position +108 to the nucleotide at position -2228, with respect to the transcription start site as set forth in sequence SEQ ID No. 1.

Again, in another embodiment, the vectors of the invention comprise the sequence encoding the ABC1 protein which is placed under the control of a
15 regulatory nucleic acid comprising the core promoter sequence and the proximal 200 bp of the ABC1 gene promoter.

As indicated above, the present invention also concerns any use of a virus as described above for preparing a pharmaceutical composition intended for treating and/or for preventing pathologies linked to cholesterol transport.

20 The present invention also concerns a pharmaceutical composition comprising one or more defective recombinant viruses as described above. These pharmaceutical compositions can be formulated with a view to topical, oral, parenteral, intranasal, intravenous, intramuscular, subcutaneous, intraocular,

transdermal, etc. administration. In one embodiment of the present invention, the pharmaceutical compositions of the invention contain a pharmaceutically acceptable vehicle for an injectable formulation, for example, for an intravenous injection, such as for example, into the portal vein of the patient. They can be, for example, 5 isotonic, sterile solutions or dry, for example, freeze-dried, compositions which, upon addition, depending on the case, of sterilized water or physiological saline, allow the constitution of injectable solutions. Direct injection into the portal vein of the patient is advantageous since it makes it possible to target the infection in the liver, and thus to concentrate the therapeutic effect on this organ.

10 The doses of defective recombinant viruses used for the injection can be adjusted as a function of various parameters, and, for example, as a function of the viral vector, the method of administration used, the pathology concerned or the desired duration of the treatment. In general, the recombinant adenoviruses according to the invention are formulated and administered in the form of doses 15 between 10^4 and 10^{14} pfu/ml, and for example, 10^6 to 10^{10} pfu/ml. The term pfu ("plaque forming unit") corresponds to the infectious power of a viral solution, and is determined by infecting a suitable cell culture and measuring, generally after 48 hours, the number of plaques of infected cells. The techniques for determining the pfu titer of a viral solution are well documented in the literature.

20 As regards retroviruses, the compositions according to the invention can directly comprise the producer cells, with a view to implanting them.

In this respect, another embodiment of the invention concerns any mammalian cell infected with one or more defective recombinant viruses as

described above. For example, the invention concerns any population of human cells infected by these viruses. They can be, for example, cells of blood origin (totipotent stem cells or precursors), fibroblasts, myoblasts, hepatocytes, keratinocytes, smooth muscle cells, endothelial cells, glial cells, etc.

5 The cells according to the invention can be derived from primary cultures. These primary cultures can be sampled by any technique known to persons skilled in the art, and then put into culture under conditions which allow their proliferation. As regards to fibroblasts, they can easily be obtained from biopsies, for example, according to the technique described by Ham [Methods Cell. Biol. 21a (1980) 255].

10 These cells can be used directly for the infection with the viruses or stored, for example, by freezing, to establish autologous banks, with a view to later use. The cells according to the invention can also be secondary cultures, obtained for example, from pre-established banks (see for example, EP 228458, EP 289034, EP 400047, EP 456640).

15 The cultured cells are then infected with the recombinant viruses, to confer upon them the capacity of producing a biologically active ABC1 protein. The infection is carried out in vitro according to techniques known to persons skilled in the art. For example, according to the type of cell used and the virus copy number desired per cell, persons skilled in the art can adjust the multiplicity of infection, and
20 optionally, the number of infection cycles carried out. It is clearly understood that the method should be performed under suitable sterile conditions when the cells are intended for an administration in vivo. The doses of recombinant virus used for the infection of the cells can be adjusted by persons skilled in the art according to the

desired aim. The conditions described above for the administration in vivo can be applied to the infection in vitro. For the infection with retroviruses, it is also possible to coculture the cells whose infection is desired with producer cells for the recombinant retroviruses according to the invention. This makes it possible to
5 dispense with purifying the retroviruses.

Another subject of the invention concerns an implant comprising mammalian cells infected with one or more defective recombinant viruses as described above or recombinant virus producer cells, and an extracellular matrix. In one embodiment of the present invention, the implants according to the invention comprise 10^5 to 10^{10}
10 cells. For example, they may comprise 10^6 to 10^8 cells.

In addition, in the implants of the invention, the extracellular matrix may comprise a gelling compound and optionally a support for anchoring the cells.

Various types of gelling agent can be used to prepare the implants according to the invention. The gelling agents may be used for embedding the cells in a matrix
15 having the constitution of a gel, and to promote the anchoring of the cells onto the support, when needed. Various cell adhesion agents can thus be used as gelling agents, such as, for example, collagen, gelatin, glycosaminoglycans, fibronectin, lectins, etc. In one embodiment of the present invention, collagen is chosen as a gelling agent. It can be collagen of human, bovine or murine origin. For example,
20 type I collagen is used.

As indicated above, the compositions according to the invention may advantageously comprise a support for anchoring the cells. The term "anchoring" refers to any form of biological and/or chemical and/or physical interaction which

leads to the adhesion and/or binding of the cells onto the support. Moreover, the cells can either cover the support used, or penetrate inside this support, or both. In the context of the invention, a nontoxic and/or biocompatible solid support may be used. For example, polytetrafluoroethylene (PTFE) fibers or a support of biological
5 origin can be used.

The present invention thus offers an effective means for treating or preventing the pathologies linked to cholesterol transport, such as obesity, hypertriglyceridemia or, in the field of cardiovascular disorders, myocardial infarction, angina, sudden death, heart failure and cerebrovascular accidents.

10 In addition, this treatment can concern both humans and any animal such as sheep, cattle, domestic animals (dogs, cats, etc.), horses, fish, etc.

RECOMBINANT HOST CELLS

The invention also concerns a recombinant host cell comprising at least one
15 of the nucleic acids of the invention chosen from sequence SEQ ID No. 1, SEQ ID No. 2, SEQ ID No. 3, SEQ ID No. 4, SEQ ID No. 5, SEQ ID No. 6, SEQ ID No. 7, and SEQ ID No. 8, for example, a nucleic acid of sequence chosen from SEQ ID No. 1, SEQ ID No. 2, SEQ ID No. 3, SEQ ID No. 4 and SEQ ID No. 5.

According to another aspect, the invention also relates to a recombinant host
20 cell comprising a recombinant vector as described above.

Host cells according to the invention are, for example, as follows:

a) prokaryotic host cells: strains of *Escherichia coli* (strain DH5- α), of *Bacillus subtilis*, of *Salmonella typhimurium*, or strains from species such as *Pseudomonas*, *Streptomyces* and *Staphylococcus*;

b) eukaryotic host cells: HeLa cells (ATCC No. CCL2), Cv 1 cells (ATCC No. CCL70), COS cells (ATCC No. CRL 1650), Sf-9 cells (ATCC No. CRL 1711), CHO cells (ATCC No. CCL-61) or 3T3 cells (ATCC No. CRL-6361), or cells of the line Hepa1-6 referenced at the American Type Culture Collection (ATCC, Rockville, MD, United States of America).

c) cells in primary culture, originating from an individual in whom the expression of a nucleic acid of interest, placed under the control of a regulatory nucleic acid according to the invention, is desired.

d) indefinitely-multiplying cells (cell lines), obtained from the cells in primary culture in c) above, according to the techniques well known to persons skilled in the art.

15

SCREENING METHODS

In vitro screening method

The invention provides methods for treating an individual affected by a pathology linked to the level of expression of the ABC1 protein. For example, such a treatment method consists in administering to the individual a compound which modifies the expression of the ABC1 gene, and which can be identified according to diverse in vitro screening methods as defined below.

A first method consists in identifying compounds which modify the expression of the ABC1 gene. According to such a method, cells expressing the ABC1 gene are incubated with a candidate substance or molecule to be tested, and the level of expression of the ABC1 messenger RNA, or the level of production of the ABC1
5 protein, is then determined.

The levels of ABC1 messenger RNA can be determined with Northern-type gel hybridization, which is well known to persons skilled in the art. The levels of ABC1 messenger RNA can also be determined by methods using PCR or the technique described by WEBB and HURSKAINEN (1996, Journal of Biomolecular
10 Screening, vol. 1 : 119).

The levels of production of the ABC1 protein can be determined by immunoprecipitation or immunochemistry, using an antibody which specifically recognizes the ABC1 protein.

According to another method for screening a candidate molecule or
15 substance which modifies the activity of a regulatory nucleic acid according to the invention, a nucleotide construct as defined above, comprising a regulatory nucleic acid according to the invention and a reporter polynucleotide placed under the control of the regulatory nucleic acid, is used, said regulatory nucleic acid comprising at least one basic promoter and at least one regulatory element from
20 one of the sequences SEQ ID No. 1 and SEQ ID No. 2, a regulatory nucleic acid comprising a region ranging from nucleotides -2228 to + 108 of sequence SEQ ID NO: 1, or a regulatory nucleic acid comprising the core promoter and the 200 pb

proximal of the ABC1 gene promoter. The reporter polynucleotide can be a gene encoding a detectable protein, such as a gene encoding a luciferase.

According to such a screening method, the cells are stably or transiently transfected with the polynucleotide construct containing the regulatory nucleic acid
5 according to the invention and the reporter polynucleotide.

The transformed cells are then incubated in the presence or absence of the candidate molecule or substance to be tested, for a sufficient time, and then the level of expression of the reporter gene is determined. The compounds which induce a statistically significant change in the expression of the reporter gene (either
10 an increase or, on the contrary, a decrease in the expression of the reporter gene) are then identified and, where appropriate, selected.

Thus, a subject of the invention is also a method for screening, in vitro, a molecule or a substance which modifies the activity of a regulatory nucleic acid according to the invention, for example, which modifies the transcription of the
15 reporter polynucleotide which is a constituent of a polynucleotide construct according to the invention, characterized in that it comprises:

- a) culturing a recombinant host cell comprising a polynucleotide of interest placed under the control of a regulatory nucleic acid according to the invention;
- 20 b) incubating the recombinant host cell with the substance or molecule to be tested;
- c) detecting the expression of the polynucleotide of interest;

d) comparing the results obtained in c) with the results obtained when the recombinant host cell is cultured in the absence of the candidate molecule or substance to be tested.

The invention also concerns a kit or pack for screening, in vitro, a candidate
5 molecule or substance which is capable of modifying the activity of a regulatory nucleic acid according to the invention, comprising:

a) a host cell transformed with a polynucleotide construct as defined above, comprising a reporter polynucleotide of interest placed under the control of a regulatory nucleic acid according to the invention; and

10 b) optionally, means for detecting the expression of the reporter polynucleotide of interest.

In one embodiment of the present invention, the reporter polynucleotide of interest is the sequence encoding a luciferase. In this embodiment, the regulatory nucleic acid according to the invention is inserted into a vector, upstream of the
15 sequence encoding the luciferase. It can be for example, the vector pGL3-basic (pGL3-b) sold by the company PROMEGA (Madison, Wisconsin, USA).

In this embodiment, the recombinant vector which comprises the sequence encoding the luciferase, placed under the control of a regulatory nucleic acid according to the invention, is transfected into hepatocellular carcinoma cells, such
20 as the cells of the line HepG2, whose luciferase activity is then determined after culturing in the presence or absence of the candidate substance or molecule to be tested.

In this embodiment, pGL3-b vectors containing either the cytomegalovirus (CMV) promoter, the ApoA1 promoter or no promoter can be used as controls. For the luciferase activity assay, the transfected cells are washed with a PBS buffer and lyzed with 500 μ l of lysis buffer (50 mM tris, 150 mM NaCl, 0.02% of sodium azide,
5 1% of NP-40, 100 μ g/ml of AEBSF and 5 μ g/ml of leupeptin).

50 μ l of the cell lysate obtained are then added to 100 μ l of the luciferase substrate (Promega) and the activity measurements are carried out on a spectrophotometric microplate reader, 5 minutes after adding the cell lysate.

The data are expressed in relative units of luciferase activity. The
10 polynucleotide constructs which produce high levels of luciferase activity in the transfected cells are those which contain a regulatory nucleic acid according to the invention, included in the sequence SEQ ID No. 1 which is capable of stimulating transcription.

For the measurements of the levels of expression of messenger RNA in a
15 screening method according to the invention, probes specific for the messenger RNA of the reporter polynucleotide of interest are first prepared, for example, using the kit multiprime labeling kit (sold by the company Amersham Life Sciences, Cleveland, Ohio, USA).

20 IN VIVO SCREENING METHOD

According to another aspect of the invention, the compositions which modify the activity of a regulatory nucleic acid according to the invention can be identified *in vivo* in non-human transgenic animals.

According to such a method, a non-human transgenic animal, for example, a mouse, is treated with a candidate molecule or substance to be tested, for example, a candidate substance or molecule which has been selected beforehand by an *in vitro* screening method as defined above.

5 After a given duration, the level of activity of the regulatory nucleic acid according to the invention is determined and compared to the activity of an identical nonhuman transgenic animal, for example, an identical transgenic mouse, which has not received the candidate molecule or substance.

10 The activity of the regulatory nucleic acid according to the invention, which is functional in the transgenic animal, can be determined by diverse methods, for example, measuring the levels of messenger RNA corresponding to the reporter polynucleotides of interest placed under the control of said regulatory nucleic acid, by Northern-type hybridization or by *in situ* hybridization.

15 According to one alternative, the activity of the regulatory nucleic acid according to the invention can be determined by measuring the levels of expression of protein encoded by the reporter polynucleotides of interest, for example, by immunohistochemistry, when the polynucleotide reporter of interest comprises an open reading frame encoding a protein which is detectable by such a technique.

20 For implementing a method for screening, *in vivo*, a candidate substance or molecule which modifies the activity of a regulatory nucleic acid according to the invention, nonhuman mammals may be used, such as mice, rats or guinea pigs or rabbits, whose genome is modified by inserting a polynucleotide construct

comprising a reporter polynucleotide of interest placed under the control of a regulatory nucleic acid according to the invention.

The transgenic animals according to the invention comprise the transgene, i.e. the above-mentioned polynucleotide construct, in a plurality of their somatic
5 and/or germ cells.

The construction of transgenic animals according to the invention can be carried out according to conventional techniques well known to persons skilled in the art. Persons skilled in the art may refer, for example, to the production of transgenic animals, such as to the production of transgenic mice, as described in Patents
10 US No. 4,873,191 (granted on 10 October 1989), US No. 5,464,764 (granted on 7 November 1995) and US 5,789,215 (granted on 4 August 1998), the contents of these documents being incorporated herein by way of reference.

Briefly, a polynucleotide construct comprising a regulatory nucleic acid according to the invention and a reporter polynucleotide of interest, placed under the
15 control of the latter, is inserted into a line of stem cells of ES type. The insertion of the polynucleotide construct is carried out for example, by electroporation, as described by Thomas et al. (1987, Cell, Vol. 51:503-512).

The cells which have undergone the electroporation are then screened for the presence of the polynucleotide construct (for example, by selection with the aid of
20 markers, or by PCR, or by Southern-type analysis of DNA on electrophoresis gels), in order to select the positive cells which have integrated the exogenous polynucleotide construct into their genome, when appropriate, following a

homologous recombination event. Such a technique is described for example, by Mansour et al. (1988, Nature, Vol. 336:348-352).

Next, the positively selected cells are isolated, cloned and injected into 3.5-day old mouse blastocysts, as is described by Bradley (1987, Production and
5 Analysis of Chimaeric mice. In: E.J. Robertson (Ed., Teratocarcinomas and embryonic stem cells: A practical approach. IRL Press, Oxford, page 113). Blastocysts are then introduced into a female host animal, and the development of the embryo is monitored until full term.

According to one alternative, positively selected cells of ES type are brought
10 into contact with 2.5-day old embryos at an 8- to 16-cell stage (*morulae*), as described by Wood et al. (1993, Proc. Natl. Acad. Sci. USA, Vol. 90:4582-4585) or by Nagy et al. (1993, Proc. Natl. Acad. Sci. USA, Vol. 90:8424-8428), the ES cells being internalized so as to extensively colonize the blastocyst, including the cells which give rise to the germ line.

15 The descendants are then tested to determine those which have integrated the polynucleotide construct (the transgene).

A subject of the invention is thus also a non-human transgenic animal whose somatic and/or germ cells have been transformed with a nucleic acid or a polynucleotide construct according to the invention.

20 The invention also relates to recombinant host cells obtained from a transgenic animal as described above.

Recombinant cell lines originating from a transgenic animal according to the invention can be established in long term culture starting from any tissue from such

a transgenic animal, for example, by transfecting primary cell cultures with vectors expressing oncogenes such as the large T antigen of SV40, as described for example, by Chou (1989, Mol. Endocrinol. Vol. 3:1511-1514) and Schay et al. (1991, Biochem. Biophys. Acta, Vol. 1072:1-7).

5 The invention also relates to a method for screening, *in vivo*, a candidate molecule or substance which modifies the activity of a regulatory nucleic acid according to the invention, comprising:

a) administering the candidate substance or molecule to a transgenic animal as described above;

10 b) detecting the level of expression of a reporter polynucleotide of interest placed under the control of the regulatory nucleic acid;

c) comparing the results obtained in b) with the results obtained in a transgenic animal which has not received the candidate substance or molecule.

The invention also relates to a kit or pack for screening, *in vivo*, a candidate
15 molecule or substance which modifies the activity of a regulatory nucleic acid according to the invention, comprising:

a) a transgenic animal as defined above;

b) optionally, the means for detecting the level of expression of the reporter polynucleotide of interest.

20

PHARMACEUTICAL COMPOSITIONS AND COMPOUNDS

The invention also concerns pharmaceutical compositions intended for preventing or for treating a deficiency in cholesterol metabolism, such as

atherosclerosis, for example, in cholesterol transport, and in the reverse transport of cholesterol.

Firstly, a subject of the invention is also a candidate substance or molecule which modifies the activity of a regulatory nucleic acid according to the invention.

5 The invention also concerns a candidate substance or molecule characterized in that it increases the activity of a regulatory nucleic acid according to the invention, and for example, of a regulatory nucleic acid comprising the sequence SED ID No.1 or SEQ ID No.3, a region comprising the sequence ranging from nucleotides -2228 to + 108 of sequence SEQ ID NO: 1, or a region comprising the
10 core promoter and the 200 pb proximal of the ABC1 gene promoter

For example, such a substance or molecule which is capable of modifying the activity of a regulatory nucleic acid according to the invention has been selected according to one of the *in vitro* or *in vivo* screening methods defined above.

Thus, an individual whose cholesterol metabolism is affected, for example, an
15 individual affected by Tangier disease, is treated by administering to this individual an effective amount of a compound which modifies the activity of a regulatory nucleic acid according to the invention.

Thus, a patient with a weak activity of the ABC1 promoter can be treated with an abovementioned molecule or substance to increase the activity of the ABC1
20 promoter.

Alternatively, a patient with an abnormally high activity of the ABC1 promoter can be treated with a compound which is capable of decreasing or blocking the activity of the ABC1 promoter.

Such a compound can be a compound which modifies the interaction of at least one transcription factor with the ABC1 promoter or a regulatory element of a regulatory nucleic acid according to the invention.

For example, the compound can inhibit the interaction of one of the
5 transcription factors listed in Table 1 with a regulatory nucleic acid according to the invention.

The compound can also be a compound which modifies the activity of a transcription factor which binds to the ABC1 promoter, or of a regulatory element present on this latter.

10 A compound of therapeutic interest according to the invention can also be a compound which modifies the interaction of a first transcription factor with a second transcription factor.

As detailed in the analysis of the various transcription factors which are capable of binding to the sequence SEQ ID No. 3, some transcription factors are
15 active only if they are associated with another transcription factor.

A compound of therapeutic interest according to the invention is for example, chosen from nucleic acids, peptides and small molecules. For example, such a compound can be an antisense nucleic acid which binds specifically to a region of the ABC1 promoter or to a regulatory element of a regulatory nucleic acid of ABC1,
20 and which inhibits or decreases the activity of the promoter.

This compound of therapeutic interest can also be an antisense nucleic acid which interacts specifically with a gene encoding a transcription factor which modifies the activity of the ABC1 promoter, such that the interaction of the antisense

nucleic acid with the gene encoding the transcription factor which binds to the ABC1 promoter decreases the production of this transcription factor, resulting in an increase or a decrease in the activity of the ABC1 promoter, according to whether the transcription factor increases or, on the contrary, reduces the activity of the
5 ABC1 promoter.

The toxicity and the therapeutic efficacy of the therapeutic compounds according to the invention can be determined according to the standard pharmaceutical protocols, in cells in culture or in experimental animals, for example, to determine the lethal dose LD₅₀ (i.e. the dose which is lethal for 50% of the
10 population tested) and the effective dose ED₅₀ (i.e. the dose which is therapeutically effective in 50% of the population tested).

For all the compounds of therapeutic interest according to the invention, the effective therapeutic dose can be estimated initially from tests carried out in cell cultures *in vitro*.

15 A subject of the invention is also pharmaceutical compositions comprising a therapeutically effective amount of a substance or molecule of therapeutic interest according to the invention.

Such pharmaceutical compositions can be formulated conventionally, using one or more physiologically acceptable vectors or excipients.

20 Thus, the compounds of therapeutic interest according to the invention, as well as physiologically acceptable salts and solvates thereof, can be formulated for administration by injection or inhalation, or by oral, buccal, parenteral or rectal administration.

Techniques for preparing pharmaceutical compositions according to the invention can be easily found by persons skilled in the art, for example, in the work Remington's Pharmaceutical Sciences, Mead Publishing Co., Easton, PA, USA.

For a systemic administration, an injection, including intramuscular, intravenous, intraperitoneal and subcutaneous injections, may be used. In this case, the pharmaceutical compositions according to the invention can be formulated in the form of liquid solutions, for example, in physiologically compatible solutions or buffers.

10 **METHOD FOR DETECTING AN IMPAIRMENT OF THE TRANSCRIPTION OF THE HUMAN ABC1 GENE**

A subject of the invention is also methods for determining whether an individual presents a risk of developing a pathology linked to a deficiency in cholesterol metabolism, such as atherosclerosis, for example, in cholesterol transport, and in the reverse transport of cholesterol, such as a risk of developing Tangier disease.

Such methods comprise detecting, in cells from a biological sample originating from an individual to be tested, the presence or absence of a genetic alteration characterized by an impairment of the expression of a gene whose expression is regulated by the ABC1 promoter.

By way of illustration, such genetic alterations can be detected, in order to determine the existence of a deletion of one or more nucleotides in the sequence of a regulatory nucleic acid for ABC1 of sequence SEQ ID No. 1 or SEQ ID No. 2, of

the addition of one or more nucleotides or of the substitution of one or more nucleotides in said sequence SEQ ID No. 1 or SEQ ID No. 2.

According to one embodiment of a method for detecting an impairment of the transcription of the ABC1 gene in an individual, the genetic alteration is identified
5 according to a method comprising sequencing all or part of the sequence SEQ ID No. 1, or alternatively all or part of at least the sequence SEQ ID No. 2.

Sequencing primers can be constructed in order to hybridize with a given region of the sequence SEQ ID No. 1. Such sequencing primers are for example, constructed so as to amplify fragments of approximately 250 to approximately 300
10 nucleotides of the sequence SEQ ID No. 1 or of a complementary sequence.

The fragments amplified, for example, by the PCR method, are then sequenced, and the sequence obtained is compared with the reference sequence SEQ ID No. 1 in order to determine whether one or more deletions, additions or substitutions of nucleotides are found in the sequence amplified from the DNA
15 contained in the biological sample originating from the individual tested.

The invention thus also concerns a method for detecting an impairment of the transcription of the ABC1 gene in an individual, comprising:

a) sequencing a nucleic acid fragment which can be amplified with the aid of at least one nucleotide primer which hybridizes with the sequence SEQ ID No. 1 or
20 SEQ ID No. 2 according to the invention;

b) aligning the sequence obtained in a) with the sequence SEQ ID No. 1 or the SEQ ID No. 2;

c) determining the presence of one or more deletions, additions or substitutions of at least one nucleotide in the sequence of the nucleic acid fragment, with respect to the reference sequence SEQ ID No. 1 or SEQ ID NO. 2.

Oligonucleotide probes which hybridize with a region of the sequence SEQ ID No. 1 or the sequence SEQ ID No. 2 in which an alteration in the sequence has been determined during the implementation of the detection method described above also form part of the invention.

Alternatively, oligonucleotide probes which hybridize specifically with a corresponding region of the sequence SEQ ID No. 1 or of the sequence SEQ ID No. 2, for which one or more deletions, additions or substitutions of at least one nucleotide have been determined in an individual, also form part of the invention.

Such oligonucleotide probes constitute means for detecting alterations in the regulatory sequence for the ABC1 gene and thus also means for detecting a predisposition to developing a pathology linked to a deficiency in cholesterol metabolism, such as atherosclerosis or Tangier disease.

A subject of the invention is thus also a kit or pack for detecting an impairment of the transcription of the ABC1 gene in an individual, comprising:

a) one or more primers which hybridize with a region of the sequence SEQ ID No. 1 or of the sequence SEQ ID No. 2;

b) optionally, the means required for carrying out an amplification reaction.

A subject of the invention is also a kit or pack for detecting an impairment of the transcription of the ABC1 gene in an individual, comprising:

a) one or more oligonucleotide probes as defined above;

b) optionally, the reagents required for carrying out a hybridization reaction.

The nucleic acid fragments derived from any one of the nucleotide sequences SEQ ID No. 1 to 8 are thus useful for detecting the presence of at least one copy of a regulatory nucleotide sequence for the ABC1 gene or of a fragment or
5 of a variant (containing a mutation or a polymorphism) of the latter, in a sample.

The nucleotide probes or primers according to the invention comprise at least 8 consecutive nucleotides of a nucleic acid chosen from the group consisting of the sequences SEQ ID No. 1 to 8, or of a nucleic acid of complementary sequence.

For example, nucleotide probes or primers according to the invention will
10 have a length chosen from 10, 12, 15, 18, 20 to 25, 35, 40, 50, 70, 80, 100, 200, 500, 1000, or 1500 consecutive nucleotides of a nucleic acid according to the invention, such as a nucleic acid of nucleotide sequence chosen from the sequences SEQ ID No. 1, SEQ ID No. 2, SEQ ID No. 3, SEQ ID No. 4, SEQ ID No. 5, SEQ ID No. 6, SEQ ID No. 7, and SEQ ID No. 8.

15 Alternatively, a nucleotide probe or primer according to the invention will consist of and/or will comprise the fragments with a length chosen from 12, 15, 18, 20, 25, 35, 40, 50, 100, 200, 500, 1000, and 1500 consecutive nucleotides of a nucleic acid according to the invention, or a nucleic acid chosen from the sequences SEQ ID No. 1, SEQ ID No. 2, SEQ ID No. 3, SEQ ID No. 4, SEQ ID No. 5, SEQ ID
20 No. 6, SEQ ID No. 7, and SEQ ID No. 8, or of a nucleic acid of complementary sequence.

The definition of a nucleotide probe and primer according to the invention thus encompasses oligonucleotides which hybridize, under the high stringency

hybridization conditions defined above, with a nucleic acid chosen from the sequences SEQ ID No. 1, SEQ ID No. 2, SEQ ID No. 3, SEQ ID No. 4, SEQ ID No. 5, SEQ ID No. 6, SEQ ID No. 7, and SEQ ID No. 8 or with a complementary sequence of these sequences.

5 Examples of primers and primer pairs for amplifying various regions of the ABC1 gene are represented below.

It is for example, the primer pair represented by the primer of sequence SEQ ID No. 12: 5'-TTG CCG TCG ACT GTT TTG GGT AGT TT-3' and the primer of sequence SEQ ID No. 13: 5'-GCC CTG TCG ACC GGC TCT GTT GGT G-3'.

10 A nucleotide primer or probe according to the invention can be prepared by any suitable method well known to persons skilled in the art, including by cloning and restriction enzyme action, or by direct chemical synthesis according to techniques such as the phosphodiester method of Narang et al. (1979) or of Brown et al. (1979), the diethylphosphoramidite method of Beaucage et al. (1980) or the
15 technique on a solid support described in EU Patent No. EP 0 707 592.

Each of the nucleic acids according to the invention, including the oligonucleotide probes and primers described above, can be labeled, if desired, by incorporating a detectable label, by spectroscopic, photochemical, biochemical, immunochemical or chemical means.

20 For example, such labels can comprise radioactive isotopes (^{32}P , ^{33}P , ^3H , ^{35}S), fluorescent molecules (5-bromodeoxyuridine, fluorescein, acetylaminofluorene, digoxigenin) or ligands such as biotin.

The labeling of the probes is for example, carried out by incorporation of labeled molecules into the polynucleotides by primer extension, or by addition onto the 5' or 3' ends.

Examples of nonradioactive labeling of nucleic acid fragments are described, for example, in French Patent No. FR 78 109 75 or in the articles by Urdea et al. (1988) or Sanchez-pescador et al. (1988).

Advantageously, the probes according to the invention can have structural properties of a type which enables an amplification of the signal, such as the probes described by Urdea et al. (1991), or in European Patent No. EP-0 225 807 (Chiron).

The oligonucleotide probes according to the invention can be used, for example, in Southern-type hybridizations to genomic DNA.

The probes according to the invention can also be used for detecting PCR amplification products or for detecting mismatches.

Nucleotide probes or primers according to the invention can be immobilized on a solid support. Such solid supports are well known to persons skilled in the art, and comprise surfaces of the wells of microtitration plates, polystyrene beds, magnetic beds, nitrocellulose bands or microparticles such as latex particles.

Consequently, the present invention also concerns a method for detecting the presence of a nucleic acid as described above in a sample, said method comprising:

1) bringing one or more nucleotide probes according to the invention into contact with the sample to be tested;

2) detecting the complex possibly formed between the probe(s) and the nucleic acid present in the sample.

According to one embodiment of the detection method according to the invention, the oligonucleotide probe(s) are immobilized on a support.

According to another aspect, the oligonucleotide probes comprise a detectable label.

5 The invention also concerns a pack or kit for detecting the presence of a nucleic acid according to the invention in a sample, said pack comprising:

- a) one or more nucleotide probes as described above;
- b) optionally, the reagents required for the hybridization reaction.

According to a first aspect, the detection pack or kit is characterized in that
10 the probe(s) is (are) immobilized on a support.

According to a second aspect, the detection pack or kit is characterized in that the oligonucleotide probes comprise a detectable label.

According to one embodiment of the detection kit described above, such a kit will comprise a plurality of oligonucleotide probes in accordance with the invention,
15 which may be used for detecting target sequences of interest, or alternatively detecting mutations in the coding regions or the noncoding regions of the nucleic acids according to the invention, for example, of the nucleic acids of sequences chosen from SEQ ID No. 1, SEQ ID No. 2, SEQ ID No. 3, SEQ ID No. 4, and SEQ ID No. 5 or the nucleic acids of complementary sequence.

20 Thus, the probes according to the invention which are immobilized on a support can be ordered in matrices such as "DNA chips". Such ordered matrices have been described, for example, in US Patent No. 5,143,854 and in PCT Applications No. WO 90/150 70 and 92/10092.

Support matrices onto which oligonucleotide probes have been immobilized at a high density are, for example, described in US Patents No. 5,412,087 and in PCT Application No. WO 95/11995.

The nucleotide primers according to the invention can be used to amplify any
5 one of the nucleic acids according to the invention, and for example, all or part of a nucleic acid of sequences chosen from SEQ ID No. 1, SEQ ID No. 2, SEQ ID No. 3, SEQ ID No. 4, and SEQ ID No. 5, , or a variant of this nucleic acid.

Another subject of the invention concerns a method for amplifying a nucleic acid according to the invention, and for example, a nucleic acid sequence chosen
10 from SEQ ID No. 1, SEQ ID No. 2, SEQ ID No. 3, SEQ ID No. 4, and SEQ ID No. 5, or a fragment or a variant of this nucleic acid, contained in a sample, said method comprising:

a) bringing the sample, in which the presence of the target nucleic acid is suspected, into contact with a pair of nucleotide primers, of which the hybridization
15 position is located respectively on the 5' side and on the 3' side of the region of the target nucleic acid whose amplification is sought, in the presence of the reagents required for the amplification reaction; and

b) detecting the amplified nucleic acids.

To implement the amplification method as described above, any one of the
20 nucleotide primers described above will advantageously be used.

A subject of the invention is also a pack or kit for amplifying a nucleic acid according to the invention, and for example, all or part of a nucleic acid sequence

chosen from SEQ ID No. 1, SEQ ID No. 2, SEQ ID No. 3, SEQ ID No. 4, and SEQ ID No. 5, said pack or kit comprising:

- a) a pair of nucleotide primers in accordance with the invention, the hybridization position of which is located respectively on the 5' side and on the 3' side of the target nucleic acid whose amplification is sought;
- b) optionally, the reagents required for the amplification reaction.

Such an amplification pack or kit will advantageously comprise at least one pair of nucleotide primers as described above.

The following examples are intended to illustrate the invention without limiting the scope thereof.

EXAMPLES:

EXAMPLE 1: Tissue distribution of the ABC1 gene transcripts according to the invention

The expression profile of the polynucleotides according to the present invention was determined according to the reverse transcription-coupled PCR and Northern blot analysis protocols described, for example, by Sambrook et al. (ref. CSH Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). "Molecular Cloning: A Laboratory Manual", 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.).

For example, in the case of an analysis by reverse transcription, a pair of primers which have been synthesized from the human ABC1 gene full-length cDNA of sequence SEQ ID No. 10 was used for detecting the corresponding cDNA.

The polymerase chain reaction (PCR) was carried out on matrices of cDNAs
5 corresponding to reverse-transcribed polyA⁺ mRNAs (Clontech). The reverse transcription into cDNA was carried out with the enzyme Superscript II (GibcoBRL, Life Technologies), according to the conditions described by the manufacturer.

The polymerase chain reaction was carried out according to standard conditions, in 20 µl of reaction mixture, with 25 ng of the cDNA preparation. The
10 reaction mixture was composed of 400 µM of each of the dNTPs, of 2 units of *Thermus aquaticus* (Taq) DNA polymerase (Ampli Taq Gold; Perkin Elmer), of 0.5 µM of each primer, of 2.5 mM MgCl₂ and of PCR buffer. Thirtyfour cycles of PCR (denaturation for 30 s at 94°C, hybridization of 30 s, broken down as follows during 34 cycles: 64°C 2 cycles, 61°C 2 cycles, 58°C 2 cycles and 55°C 28 cycles,
15 and an elongation of one minute per kilobase at 72°C) were carried out after a first denaturation at 94°C for 10 min, in a Perkin Elmer 9700 thermocycler machine. The PCR reactions were visualized on agarose gels by electrophoresis. The cDNA fragments obtained could be used as probes for a Northern blot analysis and could also be used for exactly determining the polynucleotide sequence.

20 In the case of an analysis by Northern blot, a cDNA probe produced as described above was labeled with ³²P, using the High Prime DNA labeling system (Boehringer), according to the instructions indicated by the manufacturer. After labeling, the probe was purified on a Sephadex G50 microcolumn (Pharmacia),

according to the instructions indicated by the manufacturer. The labeled and purified probe was then used for detecting the expression of the mRNAs in various tissues.

The Northern blot which contained samples of RNA from various human
5 tissues ((Multiple Tissue Northern, MTN, Clontech) Blot 2, reference 77759-1) was hybridized with the labeled probe.

The protocol which was followed for the hybridizations and washes were either directly that described by the manufacturer (Instruction manual PT1200-1), or an adaptation of this protocol using the methods which are known to persons skilled
10 in the art, and which are described for example, in F. Ausubel et al. (1999). For example, the temperatures of prehybridization and of hybridization in the presence of formamide may thus be varied.

For example, the following protocol was used:

15 **1 - Membrane competition and PREHYBRIDIZATION:**

- Mixed: 40 µl salmon sperm DNA (10 mg/ml)
 + 40 µl human placental DNA (10 mg/ml)
- Denatured for 5 min at 96°C, then immersed the mixture in ice.
- Removed the 2X SSC buffer and poured 4 ml of formamide mix into the
20 hybridization tube containing the membranes.
- Added the mixture of the two denatured DNAs.
- Incubation at 42°C for 5 to 6 hours, with rotation.

2 - Labeled probe competition:

- Added 10 to 50 μ l Cot I DNA to the labeled and purified probe, according to the amount of repeated sequences.
- Denatured for 7 to 10 min at 95°C.
- 5 - Incubated at 65°C for 2 to 5 hours.

3 - HYBRIDIZATION:

- Removed the prehybridization mix.
- Mixed 40 μ l salmon sperm DNA + 40 μ l human placental DNA; denatured 5 min at
10 96°C, then immersed in ice.
- Added 4 ml of formamide mix, the mixture of the two DNAs and the denatured labeled probe/Cot I DNA to the hybridization tube.
- Incubated for 15 to 20 hours at 42°C, with rotation.

15 4 - Washes:

- One wash at room temperature in 2X SSC to rinse.
- Twice 5 minutes at room temperature, 2X SSC and 0.1% SDS, at 65°C.
- Twice 15 minutes at 65°C, 1X SSC and 0.1% SDS, at 65°C.

After hybridization and washing, the blot was analyzed after overnight
20 exposure in contact with a phosphor screen, which was revealed with the aid of Storm (Molecular Dynamics, Sunnyvale, CA).

EXAMPLE 2: Analysis of the profile of gene expression for Tangier disease

Verification of the impairment in the expression level of the ABC1 gene, which leads to the Tangier cellular phenotype determined by hybridization of these sequences with probes corresponding to the mRNAs which originate from
5 fibroblasts from individuals possibly suffering from the disease, according to the methods described below:

1. Preparation of total RNAs, of poly(A)⁺ mRNAs and of cDNA probes

The total RNAs were obtained from cell cultures of the fibroblasts from
10 individuals who were normal or suffering from Tangier disease, by the guanidine isothiocyanate method (Chomczynski & Sacchi, 1987). The poly(A)⁺ mRNAs were obtained by affinity chromatography on cellulose-oligo(dT) columns (Sambrook et al., 1989), and the cDNAs used as probes were obtained by RT-PCR (DeRisi et al., 1997) with oligonucleotides which were labeled with a fluorescent product
15 (Amersham Pharmacia Biotech; CyDyeTM).

2. Hybridization and detection of expression levels

The glass fiber membranes containing the sequences according to the present invention which correspond to the Tangier gene were hybridized with the
20 cDNA probes, which were obtained from fibroblasts (Lyer et al., 1999). Using the Amersham/Molecular Dynamics system (Avalanche MicroscannerTM) allowed the quantification of the expressions of the sequence products in the healthy or affected cell type.

EXAMPLE 3: Use of IL-1beta-expressing THP-1 macrophages for screening molecules which activate or inhibit the expression of the ABC-1 gene.

The principle of this assay is that any substance which modifies the synthesis
5 activity of the ABC1 protein has repercussions on the synthesis of IL-1beta.

a) The macrophage cells of the THP-1 lines, which are human monocytic leukemia cells, are a model of differentiated macrophages. These cells were cultured in an RPMI 1640 medium supplemented with 10% of fetal calf serum, in multiwell plates, at a density of 2×10^5 cells per wells.

10 b) For the assay per se, the cells were then washed and placed in an RPMI 1640 medium containing 1 mg/ml of purified human albumin fraction IV.

c) The products were added into the extracellular medium. Simultaneously, the cells were then activated by addition of lipopolysaccharides (LPS) for 3 hours, at 1 µg/ml, followed by an incubation of 30 minutes in the presence of ATP at
15 5 mmol/L.

d) The concentrations of IL-1beta and of control IL-1alpha, tumor necrosis factor alpha (TNFalpha) and IL-6 were determined with ELISA kits, according to the manufacturers' instructions (R&D System; human IL-1beta Chemiluminescent ELISA reference QLB00). The variations in IL-1beta mRNA, which was not
20 supposed to be affected, were evaluated by the Northern blot technique, using the corresponding probe.

EXAMPLE 4: Expression of a gene of interest under the control of a regulatory nucleic acid for the human ABC1 gene according to the invention.

4.1 Materials and Methods

4.1.1 Construction of expression plasmids containing a regulatory nucleic acid for the human ABC1 gene.

The region of the regulatory acid for the human ABC1 gene which ranges from the nucleotide at position -995 up to the nucleotide at position +120, with respect to the transcription start site, was amplified by the PCR technique, with the aid of the following pair of primers:

10 forward primer S995 (SEQ ID No. 12), of sequence 5'-TTG CCG TCG ACT GTT TTG GGT AGT TT-3'; and

reverse primer +220R (SEQ ID No. 13), of sequence 5'-GCC CTG TCG ACC GGC TCT GTT GGT-3',

From human genomic DNA present in a BAC vector from a collection of human BAC
15 vectors.

The DNA fragment amplified was digested with the restriction endonuclease Sal 1, and then inserted into the vector PXP1 described by Nordeen et al. (1988, BioTechniques, 6: 454-457), at the Sal 1 restriction site of this vector.

The insert was then sequenced.

20

4.1.2. Cell culture and transfection

Cells of the Hepa1-6 line (ATCC, Rockville, MD, USA) were cultured in the medium E-MEM (Minimum Essential Medium with Earle's Salts), to which 10% (v/v)

of fetal calf serum (BioWhittaker, Walkersville, MD) was added.

Approximately 1.5×10^5 cells were distributed into each of the wells of a 12-well (2.5 cm) culture plate, and were cultured until approximately 50 to 70% confluence, and were then cotransformed with 1 μ g of the plasmid Sal-Lucif and
5 0.5 μ g of the control vector pBetagel (CloneTech Laboratories Inc., Palo Alto, CA, USA) using the Superfectin Reagent Kit pack (QIAGEN Inc., Valencia, CA, USA). Two hours after adding the DNA, the culture medium was removed, and replaced with complete AMEM (Minimum Essential Medium Eagle's Alpha Modification) medium.

10 After a duration of twenty hours, the cells were placed in fresh medium such as DMEM (Dulbecco's Minimum Essential Medium), to which 2 μ g/ml of glutamine, 100 units/ml of streptomycin and 0.1% of bovine serum albumin (BSA, Fraction V) were added, in the presence or absence of 50 μ g/ml of cholesterol (Sigma Chemical Co., St Louis, Missouri, MO, USA).

15 The cells were recovered 16 hours after the final change of medium using a lysis solution from the Tropix Luciferase Assay Kit pack (Tropix Inc., Bedford, MA, USA).

The cell lysate was divided into aliquot fractions which were stored at -70°C .

Freshly thawed aliquot fractions were used to quantify the proteins, using the
20 MicroBCA Kit pack (Pierce, Rockford, IL, USA), as well as to quantify the luciferase and beta-galactosidase production, using, respectively, the Tropix Luciferase Assay Kit and Galacto-Light Plus Kit packs. The assays were carried out according to the manufacturer's recommendations.

4.2 Results

The results are represented in Table 2 below:

Table 2: Luciferase production by the cells transfected with a vector containing the
5 luciferase gene placed under the control of a regulatory nucleic acid for the human
ABC1 gene according to the invention, in the presence or absence of cholesterol.

Culture conditions	Luciferase activity	β -Galactosidase activity	Normalized activity	Increase in the activity
Without cholesterol	215 523 +/- 20 018	29 548 +/- 1342	7.29	-
In the presence of cholesterol	500 126 +/- 100 069	37 741 +/- 2813	13.25	1.82

The results above show the capacity of a regulatory acid for the human ABC1
10 gene according to the invention to direct the expression of a coding sequence
placed under its control.

In addition, the regulatory nucleic acid used, which stretches from the
nucleotide at position -995 to the nucleotide at position +120, with respect to the
transcription start site, and which contains all the PPAR sites identified, was
15 regulated by cholesterol.

EXAMPLE 5: Characterization of the transcription factor binding motifs in the proximal human ABC1 gene promoter.

5.1 Materials and Methods

5 5.1.1 Construction of reporter plasmids for Luciferase Assay

Plasmids containing mutant SP1, AP1, E-box, LXR and deleted E-box fragments were constructed by site-directed mutagenesis using the overlap PCR method and the PXP1 -995 to +120 bp construct (Previato et al., *JBC*, 1991, 266:18958-63) as template. The primers listed below were used to amplify -200 to
 10 +44 bp of the human ABC1 promoter. Upper-case letters represent wild-type sequence whereas lower-case letters represent mutant sequence.

MDistal SP1F (SEQ ID NO: 14) 5'TCGCCCGTTTAgGcttgGGcgCCCGGCTC3'

MDistal SP1R: 5'GAGCCGGGGcgCCcaagCcTAAACGGGCGA3'

MProximal SP1F: (SEQ ID NO: 15) 5'CAGAGGCCGGGAgGcttgGGcgGGAGGGA3'

MProximal SP1R: 5'TCCCTCCcgCCcaagCcTCCCGGCCTCTG3'

MAP1F: (SEQ ID NO: 16) 5'CGTGCTTTCTGCTGAGgatgcGAACTAC3'

MAP1R: 5'GTAGTTCgcatcCTCAGCAGAAAGCACG3'

MEBoxF: (SEQ ID NO: 17) 5'CGGCTCCtcacggCTTTCTGCTGAGT3'

MEBoxR: 5'ACTCAGCAGAAAGccgtgaGGAGCCG3'

DEboxF: (SEQ ID NO: 18) 5'GCCTCCTTTCTGCTGAGTGA3'

DEboxR: 5'GAAAGGAGCCGGGGCCCGCCCCA3'

MLXRF: (SEQ ID NO: 19) 5'CTTTGtgtGATAGTAAActaCTGCGCTCGGTGCA

MLXRR 5'TGCACCGAGCGCAGtagTTACTATCacaCAAAG

S224-HindIII: (SEQ ID NO: 20) 5'ACTCCCAAAGCTTTGTCGTGG3'

44-HindIII: 5'GAGAAAGCTTCGGCTCGGCTCTG3'

S224-HindIII and 44-HindIII were the upstream and downstream primers used for overlap PCR. The Hind III sites are underlined. The resulting fragments which spanned -200 to +44 of the human ABC1 gene were ligated into the HindIII site of the PXP1 luciferase reporter plasmid (Nordeen et al., *Biotechniques*, 1988, 6:454-457). All constructs were confirmed by sequencing.

5.1.2. Cell culture and transfection

Murine macrophage RAW 264 7 cells and human embryonic kidney 293 cells (American Type Culture Collection, Rockville, MD) were grown in Dulbecco's modified Eagle's medium (DMEM) with 10% fetal calf serum (FCS). Approximately 1.5 x 10⁵ cells were plated in 12-well plates (Costar, Corning, NY), grown to 50-70% confluency and cotransfected with 1.5 µg of the ABC1 promoter-luciferase plasmid and 0.5 µg of the β-galactosidase vector (pCMVβ; Clontech, Palo Alto, CA) by using the Superfectin Reagent Kit (Qiagen, Valencia, CA). Three hours after addition of DNA, the cells were refed with fresh media containing 10% FCS. Sixteen hours later, cells were washed with phosphate-buffered saline (PBS) and refed with DMEM containing 0.1% bovine serum albumin (BSA) and either 50 µg/ml

cholesterol, 2 µg/ml 22-(R)-hydroxycholesterol (22(R)-Hch), 10µM 9-cis-retinoic acid (9CRA), 10-100 nM estradiol (Sigma, St. Louis, Mo), 10- 100nM regular insulin (Sigma) or 0.1% ethanol for 24 hours. After harvesting, 10 µl of cell extracts were used for luciferase and β-galactosidase assays by using the Promega dual
5 luciferase assay system (Promega, Madison, WI) or the Tropix Galacto-Light Plus Kits, respectively (Tropix, Bedford MA). The ratio of luciferase activity in relative light units was divided by the β-galactosidase activity to give a normalized luciferase value.

10 5.1.3. Gel Mobility Shift Assay

Three double-stranded human ABC1 promoter fragments (Fragment A spanning -171 to -71bp; Fragment EB spanning -156 to -130bp and Fragment EB spanning -156 to -130bp) were end-labeled with α³²P-dATP using T4 polynucleotide kinase (Lofstrand, Gaithersburg MD). Nuclear extracts were isolated from
15 unstimulated RAW 264.7 cell and HepG2 cell as well as RAW 264.7 cells after stimulation with the same concentrations of cholesterol and 22(R)-Hch shown above (Paragon Bioservices Inc. Baltimore, MD). One ng (10,000 cpm) of radiolabelled probe was added to 2.5 µg nuclear extract in 20 ul of a 20 mM TRIS gel shift buffer (pH 7.9) containing 60mM KCl, 0.2mM EDTA, 0.5mM DTT, 0.25mM PMS, 1.3mM
20 MgCl, 10% glycerol, 3% Ficoll and 3 µg of double-stranded poly (dIdC) as described (Previato et al., J Biol Chem, 1991, 266:18958-18963) and incubated for 10

minutes on ice followed by 10 minutes at room temperature. The incubated mixture was loaded on a 6% polyacrylamide gel in 0.25 x TBE buffer and electrophoresed at 100 V for 90 min followed by autoradiography. For competition assays, nuclear extracts were preincubated for 10 minutes on ice in a 20 µl reaction mixture in the presence or absence of a 100-200-fold excess of double-stranded DNA competitors for Sp1 (-173 to -155 bp), AP1 (-135 to -155 bp), LXR (-54 to -69 bp) and E-box (-158 to -136 bp) before addition of probe. For supershift assays, nuclear extracts were preincubated with antibodies against different E-box binding proteins including Mad1, Mad2, Mad3, Max, c-Myc, MyoD, USF1 and USF2 as well as Sp1, c-Fos, c-Jun, JunB and JunD (Santa Cruz Biotechnology, Santa Cruz, CA) on ice for 30 minutes before addition of probe.

5.1.4. DNase I protection Assay

End-labelled fragment A was digested with AspHI. The 94bp fragment was gel-purified from a 10% acrylamide TBE gel (Novex) and added to 14 µl of RAW cell nuclear extract (5.2 µg/µl protein) in gel shift buffer (Example 5.1.3) and incubated on ice for 10 min. One µl of probe (10,000 cpm) was added and the mixture was incubated on ice for another 10 min. After 10 min at RT, 20 µl DNaseI digestion buffer (10 mM Tris-HCl pH 8.0, 5mM CaCl₂, 5.0mM MgCl₂) was added, then 15 seconds later DNase I was added and incubated for 1 min 45 sec. DNaseI stop buffer (10 mM Tris-HCl pH 8.0, 0.6M sodium acetate pH 7.0, 0.5% SDS, 100 mM EDTA) was then added. 2.0 µl of protease K at 20 mg/ml were added and the samples were incubated at 37°C for 30 min. 10 µl of 3M NaOAc and 4.0 µl tRNA (10

mg/ml) were added. Samples were phenol/chloroform extracted and the aqueous phase was precipitated with 2.5 volumes of 100% ethanol. After a 70% ethanol wash, pellets were dissolved in sequencing gel loading buffer, heated and run on an 8% sequencing gel. Naked DNA was digested with DNaseI as described above except that nuclear extract addition and Protease K treatment were omitted. Maxam-Gilbert sequencing was performed as described in Current Protocols in Molecular Biology (Aubusel et al., *Current Protocols in Molecular Biology*, 1994, 2: 12.1-12.11).

5.1.5 Western analysis.

RAW cell nuclear extracts (35 µg protein per lane) from cells stimulated with cholesterol or 22-R-hydroxycholesterol were loaded onto NuPage Bis-Tris 4-12% gradient gels (Invitrogen, Carlsbad, CA) and run according to manufacturer's specifications. Proteins were transferred to Immobilon-P PVDF membranes (Millipore Corp, Bedford, MA). Antibodies (2 mg/ml stock) were from Santa Cruz Biotechnology, Inc. (Santa Cruz, CA) and were used according to manufacturer's specifications. Anti-USF1 C20X (catalog number sc-229X) antibody was used at a dilution of 1/600 and anti-USF2 antibodies C20X (catalog number sc862X) and N18X (catalog number sc861X) were used at a dilution of 1/1000.

5.2.1 Analysis of binding motifs in the proximal human ABC1 gene promoter

To investigate the role of some of the above described transcription factor binding motifs in this region of the human ABC1 gene promoter, *i.e.*, Sp1 (-100 and -166 bp), AP1 (-131 bp), LXR (-69bp) and E-box (-147bp), luciferase reporter

constructs under the control of the –200bp human ABC1 promoter, either wild-type (p200-L) or mutated have been generated according to Example 5.1.1. Figure 2A shows the locations of point mutations introduced into the –200 bp promoter region of the human ABC1 gene.

5 The effect of these point mutations on the transcriptional activity of the human ABC1 gene in unstimulated RAW cells is illustrated in Fig. 2B. Mutating the distal Sp1 site, the AP1 site and the proximal Sp1 sites had only a minor effect on promoter activity. In contrast, mutation of the E-box caused a strong and significant increase in promoter activity and mutation of the LXR element caused a strong and
10 significant decrease in promoter activity. These results are consistent with binding of a transcriptional repressor to the human ABC1 E-box and binding of a transcriptional activator to the LXR element.

5.2.2 Mutation of the LXR element reduces transcription of the hABC1 gene

Figure 3 shows that the LXR element is implicated as the human ABC1
15 promoter motif responsive to oxysterols. Mutation of the LXR element at –69 bp caused a significant decrease in transcription for unstimulated cells as well as for cells stimulated by cis-retinoic acid (CRA) and 22(R)-hydroxycholesterol (22OH). Mutation of the LXR element also caused a significant decrease in transcription for cholesterol-stimulated cells, presumably due to the intracellular conversion of
20 cholesterol to oxysterols. Thus, the LXR element mediates responsiveness to cholesterol as well as to cis-retinoic acid and to hydroxysterols.

5.2.3 *Mutation of the E-box increases transcription of the human ABC1 gene*

Figure 4 demonstrates that mutation or deletion of the E-box motif increases transcription of the human ABC1 gene in unstimulated RAW cells by approximately 3-fold and also in RAW cells stimulated with cholesterol (c), cis-retinoic acid (CRA) and oxysterol (22OH) by up to 40-fold. Furthermore, mutation or deletion of the E-box in the proximal human ABC1 promoter had no effect on the stimulatory effect of either CRA or oxysterols. Similar findings were demonstrated for unstimulated and stimulated human embryonal kidney 293 cells. These results indicate that the LXR-mediated activation of ABC1 gene transcription by CRA and oxysterols does not require an intact E-box motif and are consistent with binding of a transcriptional repressor to the E-box in the human ABC1 gene.

5.2.4 *Nuclear transcription factors bind the E-box in the human ABC1 gene promoter*

DNaseI footprint analysis of the ABC1 proximal promoter revealed protection of the E-box in the presence of RAW nuclear extracts (Fig. 5) indicative of a protein binding to this region. Consistent with the lack of transcriptional effects observed by independently mutating the AP1 and Sp1 motifs (Fig 2B), no protection of these potential binding sites in the human ABC1 promoter was evident by footprint analysis (Fig. 5).

To further demonstrate binding of nuclear transcription factors to the E-box motif we performed gel-shift analysis of the human ABC1 promoter (Fig. 6). The probe utilized in Figure 6B (left) was a 100 bp double-stranded fragment spanning –

171 through -71 of the ABC1 promoter (designated Fragment A). Incubation of the radiolabelled probe with nuclear extract isolated from unstimulated RAW cells resulted in a gel shift (Fig. 6B). The shift was abolished when either unlabelled Fragment A (A) or a double-stranded oligonucleotide spanning the E-box (EB) were
5 used as competitors. Competition with the 27 bp fragment encoding a scrambled E-box (EBm) did not abolish the gel shift band. Competition by double-stranded DNA fragments spanning either the proximal or distal Sp1 motifs or the AP1 binding sites did not abolish the gel shift band. Similar results were observed when nuclear extracts from unstimulated 293 cells were utilized.

10 Gel-shift analysis of the human ABC1 promoter was also performed using either a 27 bp double-stranded probe spanning the E-box (Fig. 6A; Fragment EB; right panel) or an alternative probe containing a scrambled mutant E-box sequence as described above (Fragment EBm). Incubation of the wild-type E-box probe with unstimulated RAW cell nuclear extract resulted in a gel shift, indicating binding of a
15 protein to this probe (Fig.6B; middle panel). Addition of unlabelled wild-type competitor (EB) eliminated binding. In contrast, competition with the mutant E-box fragment (EBm) did not significantly affect binding to the wild-type E-box. Moreover, using the mutant E-box as a target probe for the binding of cellular nuclear extracts did not result the formation of a gel shift band (Fig 6B; right panel). This clearly
20 demonstrates specific binding of a protein to the wild-type E-box motif of the human ABC1 gene.

5.2.5 *USF binds the E-box in the hABC1 gene promoter*

In the human ABC1 promoter, the E-box is flanked by two C's, leading to a sequence of CCACGTGC. This is a perfect match to the consensus motif for the transcription factor USF. To establish that USF is in fact the transcription factor that
5 that binds to the E-box in the hABC1 gene promoter, a gel shift analysis utilizing USF-specific antibodies was performed (Fig. 6C). Using the 100-bp fragment as a probe (Fig. 6C, left panel), it was demonstrated that the E-box gel-shifts some protein in the RAW cell nuclear extract. Addition of anti-USF antibodies against either the amino (N) or carboxy (C) terminus of USF1 or USF2 caused a supershift
10 of the gel-shifted probe, confirming the identity of the E-box binding proteins as USF1 and USF2. Antibodies against other E-box binding proteins including Mad1, Mad2, Mad3, Max, c-Myc, and MyoD as well as Sp1, c-Jun, JunB and JunD did not compete or supershift the DNA-protein gel shift band.

Similar results were obtained by preincubating the 27 bp double-stranded
15 fragment spanning the E-box (EB) with anti-USF antibodies (Fig. 6C; right panel). As with the 100bp gel shift fragment, antibodies specific to other members of the helix-loop-helix family of transcription factors known to also bind the E-box motif did not alter the gel shift band obtained with the 27bp EB probe. No differences have been shown in the gel shift banding patterns obtained when nuclear extracts isolated from
20 unstimulated RAW cells and RAW cells stimulated with cholesterol or oxysterols were incubated with the EB probe.

This combined data identify USF1 and USF2 as the transcriptional factors that bind to the E-box in the proximal ABC1 promoter and indicate that their binding is not modulated by known activators of ABC1 gene expression.

5

5.2.6 USF1 and USF2 are expressed in RAW cells

The presence of USF in RAW cell nuclear extracts was established by performing Western blot hybridization analyses utilizing antibodies specific to USF1 (N- and C- terminus) and USF2 (N- and C-terminus). Two major immunoreactive
10 bands of approximately 43 and 44 KDa in size were identified whose expression in RAW cells were not altered by stimulation with cholesterol or oxysterols. Importantly, expression of the 18 kDa mini-USF isoform which lacks the carboxy-terminus transcriptional activating domain (Liu et al., JBC, 1999, 274: 35037-35045; Sirito et al., Nucl Acids Res, 1994, 22:427-433) was not detected in either
15 unstimulated RAW cells or RAW cells incubated for 24 hours with cholesterol or oxysterols.

All articles, patents, and patent applications mentioned in this specification are herein incorporated by reference.

20

REFERENCES

- Altschul S.F. et al., J. Mol. Biol. 1990 215: 403-410.
- Altschul S.F. et al., Nucleic Acids Res. 1997 25 : 3389-3402.
- 5 • Apfel et al., 1994, Moll. Cell. Biol. 14: 7025-7035.
- Ausubel et al., 1999, Current Protocols in Molecular Biology, Green Publishing Associates and Wiley Interscience, N.Y.
- Beard et al., Virology 75 (1990) 81.
- Beaucage et al., *Tetrahedron Lett* 1981, 22: 1859-1862.
- 10 • Bender et al., J. Virol. 61 (1987) 1639.
- Bernstein et al., Genet. Eng. 7 (1985) 235.
- Bodzioch M. et al., Nat Genet. 1999, 22: 347-351.
- Bradley, 1987, Production and analysis of Chimaeric Mice. In: E.J. Robertson Ed., Teratocarcinomas and embryonic stem cells; A practical approach. IRL Press, Oxford, page 113.
- 15 • Breakfield et al., New Biologist 3 (1991) 203.
- Brooks-Wilson et al., Nat Genet, 1999, 22: 336-345.
- Brown EL. Belagaje R, Ryan MJ, Khorana HG, *Methods Enzymol* 1979; 68: 109-151.
- 20 • Catala et al., 1989, Nucleic Acids Research, Vol. 17: 3811-3827.
- Chen et al., 1987, Mol. Cell. Biol., 7: 2745-2752.
- Chomczynski, P. and Sacchi, 1987, Anal. Biochem., 162, 156-159.

- **Chou, 1989**, Mol. Endocrinol., Vol. 3: 1511-1514.
- **DeRisi J. et al.**, 1997, Science, 278, 680-686.
- **Felgner et al.**, PNAS 84 (1987) 7413.
- **Fikes et al.**, 1990, Nature, Vol. 346: 291-294.
- 5 • **Flotte et al.**, 1992, *Am. J. Respir. Cell Mol. Biol.*, 7: 349-356.
- **Fraley et al.**, 1979, Proc. Natl. Acad. Sci. USA, 76: 3348-3352.
- **Fraley et al.**, J. Biol. Chem. 255 (1980) 10431.
- **Fuller S.A. et al.**, 1996, *Immunology in Current Protocols in Molecular Biology*, Ausubel et al.
- 10 • **Gill et al.**, 1991, Cell, Vol. 65: 333-340.
- **Gopal**, 1985, Mol. Cell. Biol., 5: 1188-1190.
- **Graham et al.**, 1973, Virology, 52: 456-457.
- **Graham et al.**, J. Gen. Virol. 36 (1977) 59.
- **Graham**, 1984, EMBO J., Vol. 3: 2917.
- 15 • **Grange et al.**, 1991, Nucleic Acids Research, Vol. 19: 131-139 Ham, Methods Cell. Biol. 21a (1980) 255.
- **Hames BD and Higgins SJ**, 1985, "Nucleic acid hybridization: a practical approach", Hames and Higgins Ed., IRL Press, Oxford.
- **Harland et al.**, 1985, J. Cell. Biol., 101: 1094-1095.
- 20 • **Hoffmann et al.**, 1990, Gene Dev., Vol. 4: 1141-1148.
- **Huygen et al.**, 1996, Nature Medicine, 2(8): 893-898.
- **Janowski et al.**, 1999, Proc. Natl. Acad. Sci. USA, Vol. 96: 266-271.

- Kaneda et al., Science 243 (1989) 375.
- Kim et al., 1990, Mol Cell Biol, Vol. 10: 5958-5966
- Langmann T. et al., 1999, Biochem. Biophys. Res. Comm., 257: 29-33.
- Levrero et al., Gene 101 (1991)
- 5 • Lim et al., 1992, J. Biol. Chem., Vol. 268: 18008-18017.
- Luciani M.F. et al., 1994, Genomics, 21: 150-159.
- Lum et al., 1990, Mol Cell Biol, Vol. 10: 6709-6717.
- Lyer V. et al., 1999, Science, 283: 83-87.
- Mansour et al., 1988, Nature, Vol. 336: 348-352.
- 10 • Mantovani et al., 1988, Nucleic Acids Research, Vol. 16: 4299-4313.
- Marcil M. et al., 1999, The Lancet, Vol. 354: 1341-1346
- McCormick, BioTechnology 3 (1985) 689.
- McLaughlin BA et al., 1996, Am. J. Hum. Genet., 59: 561-569.
- Nagy et al., 1993, Proc. Natl. Acad. Sci. USA, Vol. 90: 8424-8428.
- 15 • Narang SA, Hsiung HM, Brousseau R, Methods Enzymol 1979; 68: 90-98.
- Nicolau C. et al., 1987, Methods Enzymol., 149: 157-76.
- Overdier et al., 1994, Mol Cell Biol, Vol. 14: 2755-2766.
- Pagano et al., J. Virol. 1 (1967) 891.
- Remaley A. et al., Proc Nat Acad Sci USA, 1999, 96: 12685-12690.
- 20 • Remington's Pharmaceutical Sciences, Mead Publishing Co., Easton, PA.
- Rosenthal et al., 1990, Nucleic Acids Research, Vol. 18: 6239.
- Rust S. et al., Nature Genetics, Vol. 20, September 1998, pages 96-98.

- Rust S. et al., Nature Genetics, Vol. 22, August 1999, pages 352-355.
- Sambrook, J. Fritsch, E.F., and T. Maniatis, 1989. Molecular cloning: a laboratory manual. 2ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
- 5 • Samulski et al., 1989, *J. Virol.*, 63: 3822-3828.
- Sanchez-Pescador R., 1988, *J. Clin. Microbiol.*, 26 (10): 1934-1938.
- Schay et al., 1991, *Biochem Biophys Acta*, Vol. 1072: 1-7.
- Song et al., 1994, *Proc. Natl. Acad. Sci. USA*, Vol. 91: 10809-10813.
- Sternberg N.L., 1992, *Trends Genet.*, 8: 1-16.
- 10 • Sternberg N.L., 1994, *Mamm. Genome*, 5: 397-404.
- Tacson et al., 1996, *Nature Medicine*, 2 (8): 888-892.
- Thomas et al., 1987, *Cell*, Vol. 51: 503-512.
- Tur-Kaspa et al., 1986, *Mol. Cell. Biol.*, 6: 716-718.
- Urdea M.S., 1988, *Nucleic Acids Research*, 11: 4937-4957.
- 15 • Urdea M.S. et al., 1991, *Nucleic Acids Symp Ser.*, 24: 197-200.
- Wagner et al., 1990, *EMBO J.*, Vol. 9: 4477-4784.
- Wang et al., 1993, *Mol. Cell Biol.* Vol. 13: 5691-5701.
- Webb and Hurskainen, 1996, *Journal of Biomolecular Screening*, Vol. 1: 119.
- Willy et al., 1995, *Genes Dev.* 9: 1033-1045.
- 20 • Wood et al., 1993, *Proc. Natl. Acad. Sci. USA*, Vol. 90: 4582-4585.

WHAT IS CLAIMED IS:

1. An isolated nucleic acid comprising a polynucleotide having at least 20 consecutive nucleotides of the nucleotide sequence SEQ ID No. 1, or an isolated nucleic acid of complementary sequence.

5 2. The isolated nucleic acid according to claim 1, comprising a polynucleotide which has at least 20 consecutive nucleotides of the sequence SEQ ID No. 2, or an isolated nucleic acid of complementary sequence.

3. The isolated nucleic acid according to claim 1, comprising a polynucleotide which has at least 20 consecutive nucleotides of the sequence SEQ ID No. 3, or an
.0 isolated nucleic acid of complementary sequence.

4. The isolated nucleic acid according to claim 1, comprising a polynucleotide which has at least 20 consecutive nucleotides of the sequence SEQ ID No. 4, or an isolated nucleic acid of complementary sequence.

5. The isolated nucleic acid according to claim 1, comprising a polynucleotide
1.5 which has at least 20 consecutive nucleotides of the sequence SEQ ID No. 5, or an isolated nucleic acid of complementary sequence.

6. The isolated nucleic acid according to claim 1, wherein said nucleic acid modifies the transcription of a polynucleotide placed under its control.

7. The isolated nucleic acid according to claim 6, wherein said isolated nucleic
2.0 acid is a polynucleotide comprising a sequence ranging from the nucleotide at position -1

to the nucleotide at position -200, with respect to the first nucleotide transcribed, which is located at position 2894 of the nucleotide sequence SEQ ID No. 1.

8. The isolated nucleic acid according to claim 6, wherein said isolated nucleic acid is a polynucleotide comprising a sequence ranging from the nucleotide at position -1 to the nucleotide at position -300, with respect to the first nucleotide transcribed, which is located at position 2894 of the nucleotide sequence SEQ ID No. 1.

9. The isolated nucleic acid according to claim 6, comprising a polynucleotide ranging from the nucleotide at position -1 to the nucleotide at position -600, with respect to the first nucleotide transcribed, which is located at position 2894 of the nucleotide sequence SEQ ID No. 1.

10. The isolated nucleic acid according to claim 6, comprising a polynucleotide ranging from the nucleotide at position -1 to the nucleotide at position -2894, with respect to the first nucleotide transcribed, which is located at position 2894 of the nucleotide sequence SEQ ID No. 1.

11. The isolated nucleic acid according to claim 6, comprising a polynucleotide ranging from the nucleotide at position +120 to the nucleotide at position -995, with respect to the first nucleotide transcribed, which is located at position 2894 of the nucleotide sequence SEQ ID No. 1.

12. The isolated nucleic acid according to claim 6, comprising a polynucleotide ranging from the nucleotide at position +108 to the nucleotide at position -2228, with

respect to the first nucleotide transcribed, which is located at position 2894 of the nucleotide sequence SEQ ID No. 1.

13. The isolated nucleic acid according to claim 6, wherein said isolated nucleic acid activates the transcription of a polynucleotide of interest placed under its control.

5 14. The isolated nucleic acid according to claim 6, wherein said isolated nucleic acid inhibits the transcription of a polynucleotide of interest placed under its control.

15. An isolated nucleic acid having at least 80% nucleotide identity with an isolated nucleic acid according to claim 1.

10 16. The isolated nucleic acid according to claim 15, wherein said isolated nucleic acid modifies the transcription of a polynucleotide placed under its control.

17. The isolated nucleic acid according to claim 15, wherein said isolated nucleic acid is a polynucleotide comprising a sequence ranging from the nucleotide at position -1 to the nucleotide at position -300, with respect to the first nucleotide transcribed, which is located at position 2894 of the nucleotide sequence SEQ ID No. 1.

15 18. The isolated nucleic acid according to claim 15, comprising a polynucleotide ranging from the nucleotide at position -1 to the nucleotide at position -600, with respect to the first nucleotide transcribed, which is located at position 2894 of the nucleotide sequence SEQ ID No. 1.

20 19. The isolated nucleic acid according to claim 15, comprising a polynucleotide ranging from the nucleotide at position -1 to the nucleotide at position -200, with respect to

the first nucleotide transcribed, which is located at position 2894 of the nucleotide sequence SEQ ID No. 1.

20. The isolated nucleic acid according to claim 15, comprising a polynucleotide ranging from the nucleotide at position -1 to the nucleotide at position -2894, with respect
5 to the first nucleotide transcribed, which is located at position 2894 of the nucleotide sequence SEQ ID No. 1.

21. The isolated nucleic acid according to claim 15, comprising a polynucleotide ranging from the nucleotide at position +120 to the nucleotide at position -995, with respect
10 to the first nucleotide transcribed, which is located at position 2894 of the nucleotide sequence SEQ ID No. 1.

22. The isolated nucleic acid according to claim 15, comprising a polynucleotide ranging from the nucleotide at position +108 to the nucleotide at position -2228, with respect to the first nucleotide transcribed, which is located at position 2894 of the nucleotide sequence SEQ ID No. 1.

15
23. An isolated nucleic acid which hybridizes, under high stringency hybridization conditions, with an isolated nucleic acid according to claim 1.

24. The isolated nucleic acid according to claim 23, wherein said isolated nucleic acid modifies the transcription of a polynucleotide placed under its control.

25. The isolated nucleic acid according to claim 24, wherein said isolated nucleic acid is a polynucleotide comprising a sequence ranging from the nucleotide at position -1 to the nucleotide at position -300, with respect to the first nucleotide transcribed, which is located at position 2894 of the nucleotide sequence SEQ ID No. 1.

5 26. The isolated nucleic acid according to claim 24, wherein said isolated nucleic acid is a polynucleotide comprising a sequence ranging from the nucleotide at position -1 to the nucleotide at position -200, with respect to the first nucleotide transcribed, which is located at position 2894 of the nucleotide sequence SEQ ID No. 1.

27. The isolated nucleic acid according to claim 24, comprising a polynucleotide
10 ranging from the nucleotide at position -1 to the nucleotide at position -600, with respect to the first nucleotide transcribed, which is located at position 2894 of the nucleotide sequence SEQ ID No. 1.

28. The isolated nucleic acid according to claim 24, comprising a polynucleotide
15 ranging from the nucleotide at position -1 to the nucleotide at position -2894, with respect to the first nucleotide transcribed, which is located at position 2894 of the nucleotide sequence SEQ ID No. 1.

29. The isolated nucleic acid according to claim 24, comprising a polynucleotide
20 ranging from the nucleotide at position +120 to the nucleotide at position -995, with respect to the first nucleotide transcribed, which is located at position 2894 of the nucleotide sequence SEQ ID No. 1.

30. The isolated nucleic acid according to claim 24, comprising a polynucleotide ranging from the nucleotide at position +108 to the nucleotide at position -2228, with respect to the first nucleotide transcribed, which is located at position 2894 of the nucleotide sequence SEQ ID No. 1.

5 31. The isolated nucleic acid according to claim 24, wherein said isolated nucleic acid activates the transcription of a polynucleotide of interest placed under its control.

32. The isolated nucleic acid according to claim 24, wherein said isolated nucleic acid inhibits the transcription of a polynucleotide of interest placed under its control.

10 33. An isolated nucleic acid comprising the isolated nucleic acid according to one of claims 1 to 32, further comprising a polynucleotide encoding at least one compound chosen from polypeptides of interest and nucleic acids of interest.

34. The isolated nucleic acid according to claim 33, wherein said polynucleotide encoding at least one compound encodes at least one nucleic acid of interest chosen from sense oligonucleotides and antisense oligonucleotides.

15 35. A recombinant vector comprising at least one isolated nucleic acid according to one of claims 1 to 34.

36. The recombinant vector according to claim 35, wherein said vector is chosen from a recombinant cloning vector and a recombinant expression vector.

20 37. A host cell transformed with at least one isolated nucleic acid according to one of claims 1 to 34.

38. A host cell transformed with a recombinant vector according to claim 35.

39. A non-human transgenic mammal wherein at least one of said mammals cells chosen from somatic cells and germ cells have been transformed with at least one isolated nucleic acid according to one of claims 1 to 34.

5 40. A non-human transgenic mammal wherein at least one of said mammals cells chosen from somatic cells and germ cells have been transformed with a recombinant vector according to claim 35.

41. A method for screening a substance or a molecule which modifies the transcription of the polynucleotide which is a constituent of the isolated nucleic acid
0 according to claim 33, comprising:

- a) culturing a host cell transformed according to one of claims 37 and 38;
- b) incubating the transformed host cell in the presence of the candidate substance
or molecule;
- c) detecting the expression of the polynucleotide of interest;
- 5 d) comparing the detection results obtained in c) with the detection results obtained
by culturing the transformed host cell in the absence of the candidate molecule or
substance.

42. A kit for screening, *in vitro*, a candidate molecule or substance which modifies the transcription of the polypeptide of interest which is a constituent of the isolated nucleic
10 acid according to claim 33, comprising:

- a) a host cell transformed according to one of claims 37 and 38;

b) optionally, the means required for detecting the transcription of the polynucleotide of interest which is a constituent of the isolated nucleic acid according to claim 30.

43. A method for screening, *in vivo*, a substance or molecule which modifies the transcription of a polynucleotide of interest which is a constituent of the isolated nucleic acid according to claim 33, comprising:

a) administering the substance or molecule to a nonhuman transgenic mammal according to one of claims 39 and 40;

b) detecting the expression of the polynucleotide of interest in the transgenic mammal as treated in a);

.0 c) comparing the detection results in b) with the results observed in a nonhuman transgenic mammal according to one of claims 39 and 40 which has not received the administration of the candidate substance or molecule.

44. A kit or pack for screening, *in vivo*, at least one candidate molecule or substance which modifies the transcription of the polynucleotide of interest which is a constituent of the isolated nucleic acid according to one of claims 1, 15 and 23, comprising:

a) a non-human transgenic mammal according to one of claims 39 and 40;

b) optionally, the means required for detecting the transcription of said polynucleotide of interest.

45. A substance which modifies the transcription of a polynucleotide of interest which is a constituent of the isolated nucleic acid according to claim 33.

46. The substance according to claim 45, wherein said substance comprises at least one molecule which modifies the transcription of a polynucleotide of interest which is a constituent of the isolated nucleic acid according to claim 33.

47. The substance according to claim 45, wherein said substance is selected
5 according to the method of claim 41 or the method of claim 43.

48. A pharmaceutical composition comprising, as active principle, at least one substance or a molecule according to one of claims 45 to 47.

49. The pharmaceutical composition according to claim 48, wherein said composition is used for treating a disorder chosen from hypercholesterolemia and
0 atherosclerosis.

50. A substance according to one of claims 45 to 47, as an active principle of a medicinal product.

51. A method for detecting an impairment of the transcription of the ABC1 gene in an individual, comprising:

- 5 a) extracting the total messenger RNA from a biological material originating from the individual to be tested;
- b) quantifying the ABC1 messenger RNA present in said biological material;
- c) comparing the amount of ABC1 messenger RNA obtained in b) with the amount of ABC1 messenger RNA expected in a normal individual.

52. A method for detecting an impairment of the transcription of the ABC1 gene in an individual, comprising:

a) sequencing, starting with a biological material originating from said individual, a polynucleotide located upstream of the transcription start site of the ABC1 gene;

5 b) aligning the nucleotide sequence obtained in a) with the sequence SED ID No. 1;

c) determining the nucleotide differences between the sequenced polynucleotide originating from the biological material of the individual to be tested and the reference sequence SEQ ID No. 1.

53. A kit for detecting an impairment of the transcription of the ABC1 gene in an individual, comprising the means required for quantifying the ABC1 messenger RNA in a biological material originating from said individual.

54. A kit for detecting an impairment of the transcription of the ABC1 gene in an individual, comprising the means required for sequencing a polynucleotide located upstream of the transcription start site of the ABC1 gene in said individual.

15 55. A method for screening a molecule or substance which modifies the transcription of the polynucleotide of interest which is a constituent of an isolated nucleic acid according to claim 33, comprising:

a) incubating at least one isolated nucleic acid according to one of claims 1 to 34, or a recombinant vector according to claim 35, with a candidate molecule or substance to be
20 tested;

b) detecting the complex formed between the candidate molecule or substance and the candidate molecule or substance.

56. A kit or pack for screening a candidate molecule or substance which modifies the transcription of the polynucleotide of interest which is a constituent of the isolated nucleic acid according to claim 33, comprising:

a) at least one isolated nucleic acid according to one of claims 1 to 34 or a recombinant vector according to claim 35;

b) optionally, the means required for detecting the complex formed between the candidate molecule or substance and said isolated nucleic acid.

SEQUENCES LISTING

<110> AVENTIS PHARMA S.A.

<120> Regulatory nucleic acid for the ABC1 gene, molecules modifying its activity, and therapeutic uses

<130> ABC1 Promoter - Aventis

<140>

<141>

<160> 20

<170> PatentIn Ver. 2.1

<210> 1

<211> 3231

<212> DNA

<213> Homo sapiens

<400> 1

acagggcatg	gtggcaggtg	cctgtaatct	cagttactcg	ggaggtggag	gttgcaatga	60
gccagatcg	caccattgca	ctccagcctg	ggcaacaaaa	ggtgaaactc	catctcaatt	120
aaaaaaaaaa	gaatgatttt	ggtgggtcgac	ttcaaataagg	taggagaaga	aggagagagg	180
agatggagg	tcagggagat	ctaattactc	tctaaaatca	tgctaggaaa	gataacacct	240
tttaataaca	ctctctgctt	ttataacatc	attctgccaa	ggagctcaaa	ggtttcaaca	300
aagttcactt	tcagaaaacc	cctttgagga	agacagaata	tacatcttct	ctccatttta	360
aagatgaaga	aacaggccgg	gcacaatggc	taatgcctgt	aatcccagca	ctttgggagg	420
ctgaggccag	aggatcgctt	gagctccaga	gtttgagacc	agcctggata	acatggcaaa	480
accctgtctc	tacaaaaaaa	atacaaaaaa	tagatgggtg	tggtggcatg	cacctgtggt	540
cccagctact	tgggaggcta	aggtgggagg	atcgcttgag	cccagggagt	caagtctaca	600
ctgagccatg	attggatcac	tgcactccag	cctgggtaga	cagagcaaga	ccctgtctca	660
aaaaaaagaa	atgaaagaga	aagaaagaaa	gaggagagga	gaggagatga	ggggaggagg	720
gaggggggga	ggaaggaagg	aaggaaggaa	ggaaaaaaag	atgaaaaaag	aaaaaaacaa	780
gatgaaacag	aggcagaaag	actttacgta	aattgtctcat	catgtgggtt	tcaagtttga	840
ccccaaaacc	caatttattg	accaaggtta	ttctttgact	gaggcaaggg	ggtccgctct	900
cctgggcctt	gggctttaga	aagctcatct	ctggcctttc	tgagatccat	ccctttcttt	960
ttatttttct	tgacacggag	tcttgctctg	tcactcaggc	tggagtgcag	tggcatgatc	1020
tcgactcact	gtaacctctg	cctcccgggt	tcaagcgatt	ctcctgcctc	agcctcctga	1080
gataacaggc	gcccgccacc	acatctggct	aatttttgta	tttttagtaa	agactggggt	1140
tcacatgttt	ggccagggtt	gttttcgaact	cctgacctga	ggtgagctgc	ccaccttggc	1200
ctcccaaagt	gctgggatta	caggcatgag	ccactgcgcc	cagctcagat	ccatcccttt	1260
ctaagggcaa	acagtccatg	gtgcaaaggg	gccatgccac	ccagagttaa	gagtacctgg	1320
gactccagaa	ttccttgcc	ggtggcctcc	acatgcactt	ccagggcctg	cttgggcctc	1380
ttctatgcgt	ctgtcctgag	tgttgataga	accactgatg	tgagtacctg	ggcttgagcc	1440
gtggcctgga	gactcctgtt	actgtagcat	ggagggggct	tgtgcagctg	aatgtctgca	1500
tgcaggtggt	gggagttctg	gaatatgatg	gagctggagg	tgggaagaga	agtaggcttg	1560
gggcagctct	ctcatgccac	ctcattcttg	ccaaaactca	ggtcaaactg	tgaagagtct	1620
aaatgtgaat	ctgcccttca	aggtggctac	aaaggtatct	ttgtcaaggt	aggagacctt	1680
gtggcctcca	cgtgcacttc	cagggcctgc	ttgggcctct	tctacgggtc	tgtcctgagt	1740
cttctatgaa	tccttcaggg	cagattcata	tttagactct	tcacagtttg	acctgagttt	1800
tggccagaat	aaggtgacat	ttagtttggt	ggcttgatgg	atgacttaaa	tatttagaca	1860
tgggtgtgat	gcctgcattc	ctactcttgc	cttttttttt	gcccctccag	tgttttgggt	1920
agttttgtct	ccctacagcc	aaaggcaaac	agagaagttg	gaggtctgga	gtggctacat	1980
aattttacac	gactgcaatt	ctctggctgc	acttcacaaa	tgtatacaaa	ctaaatacaa	2040
gtcctgtgtt	tttatcacag	ggaggctgat	caatataatg	aaattaaaaa	ggggctgggtc	2100
catattgttc	tgtgtttttg	ttgttttggt	ttgttttggt	ctttttttgt	ttttgtggcc	2160
tccttctctc	caatttatga	agagaagcag	taagatgttc	ctctcgggtc	ctctgaggga	2220
cctggggagc	tcaggctggg	aatctccaag	gcagtaggtc	gcctatcaaa	aatcaaagtc	2280
cagggttttg	gggggaaaac	aaaagcagcc	cattaccagc	aggactgtcc	gccttcccct	2340
caccccagcc	taggcctttg	aaaggaaaaca	aaagacaaga	caaaatgatt	ggcgctcctga	2400
gggagattca	gcctagagct	ctctctcccc	caatccctcc	ctccgggtga	ggaaactaac	2460
aaaggaaaaa	aaaattgcgg	aaagcaggat	ttagaggaag	caaattccac	tgggtgccctt	2520
ggctgcgggg	aacgtgggact	agagagtctg	cggcgagcgc	ccgagcccag	cgcttcccgc	2580
gcgtcttagg	ccggcggggc	cgggcggggg	aaggggacgc	agaccgcgga	ccctaagaca	2640
cctgctgtac	cctccacccc	caccccaccc	cacccacctc	cccccaactc	cctagatgtg	2700
tcgtggggcg	ctgaacgtcg	ccggtttaag	gggcgggccc	cggctccacg	tgctttctgc	2760
tgagtgactg	aactacataa	acagaggccg	ggaagggggc	ggggaggagg	gagagcacag	2820

```
caggtttgtg gggggaaaac aaaagcagcc cattaccag aggactgtcc gccttccccct 2340
caccaccagcc taggcctttg aaaggaaaca aaagacaaga caaatgatt ggcgtcctga 2400
gggagattca gcctagagct ctctctcccc caatccctcc ctccggctga ggaaactaac 2460
aaaggaaaaa aaaattgcgg aaagcaggat ttagaggaag caaattccac tgggtgccctt 2520
ggctgccggg aacgtggact agagagtctg cggcgagcc cagagccag cgcttccccg 2580
gcgtcttagg ccggcggggc cgggcggggg aaggggacgc agaccgcgga ccctaagaca 2640
cctgctgtac cctccacccc caccaccacc caccacctc cccccaactc cctagatgtg 2700
tcgtgggcgg ctgaacgtcg cccgtttaag gggcgggccc cggctccacg tgctttctgc 2760
tgagtgactg aactacataa acagaggccg ggaagggggc ggggaggagg gagagcacag 2820
gctttgaccg atagtaacct ctgcgctcgg tgcagccga tctataaaag gaactagtcc 2880
cggcaaaaac cccgtaattg cgagcgagag tgagtggggc cgggacccgc agagccgagc 2940
cgacccttct ctcccgggct gcggcagggc agggcgggga gctccgcgca ccaacagagc 3000
cggttctcag ggcgctttgc tccttgtttt tccccgggt ctgttttctc cccttctcgg 3060
gaaggcttgt caaggggtag gagaaagaga cgcaaacaca aaagtggaaa acaggttaaga 3120
ggctctccag tgacttactt gggcgttatt gttttgtttc gaggccaagg aggccttcggg 3180
aagtgtcggg tttcggggac tttgatccgg agccccacat cccaccact t 3231
```

<210> 2

<211> 357

<212> DNA

<213> Homo sapiens

<400> 2

```
tggaggtctc agctgagagg gctggattag cagtctcat tgggtgatgg ctttgcagca 60
ataactgatg gctgtttccc ctctgcttt atctttcagt taatgaccag ccacgggctg 120
ccctgctgtc agctctggcc gctgccttcc agggctccc agccacacgc tgggctgtgt 180
ggctgagggg acatggcatg ttggcctcag ctgaggttgc tgctgtggaa gaacctcact 240
ttcagaagaa gacaaacagt aagcttgggt ttttcagcag cgggggggtc tctcattttt 300
tctttgtggt tttgagttgg ggattggagg agggaggagg ggaaggaagc tgtgttg 357
```

<210> 3

<211> 2893

<212> DNA

<213> Homo sapiens

<400> 3

```
acagggcatg gtggcaggtg cctgtaatct cagttactcg ggaggtggag gttgcaatga 60
gccagatcg caccattgca ctccagcctg ggcaacaaaa ggtgaaactc catctcaatt 120
aaaaaaaaa gaatgatttt ggtggctgac ttcaaatagg taggagaaga aggagagagg 180
agatggaggg tcagggagat ctaattactc tctaaaatca tgctaggaaa gataacacct 240
tttaataaca ctctctgctt ttataacatc attctgccaa ggagctcaaa ggtttcaaca 300
aagttcactt tcagaaaacc cctttgagga agacagaata tacatcttct ctccatttta 360
aagatgaaga aacaggccgg gcacaatggc taatgcctgt aatcccagca ctttgggagg 420
ctgaggccag aggatcgctt gagctccaga gtttgagacc agcctggata acatggcaaa 480
accctgtctc tacaaaaaaa atacaaaaat tagatgggtg tgggtggcatg cacctgtggt 540
cccagctact tgggaggcta aggtgggagg atcgcttgag cccagggagt caagtctaca 600
ctgagccatg attggatcac tgcaactccag cctgggtaga cagagcaaga ccctgtctca 660
aaaaaaagaa atgaaagaga aagaaagaaa gaggagagga gaggagatga ggggaggagg 720
gagggggggg ggaagggaagg aaggaaggaa ggaaaaaag atgaaaaaag aaaaaaaca 780
gatgaaacag aggcagaaag actttacgta aattgtcat catgtgggtt tcaagtttga 840
ccccaaaacc caattttatt accaaggtta ttctttgact gaggcaaggg ggtccgctct 900
cctgggcctt gggctttaga aagctcatct ctggcctttc tgagatccat cccttcttt 960
ttatttttct tgacacggag tcttgctctg tcaactcagg tggagtgcag tggcatgatc 1020
tgactcact gtaacctctg cctcccggtg tcaagcgatt ctctgcctc agcctcctga 1080
gataacaggc gcccgccacc acatctggct aatttttgta tttttagtaa agactgggtt 1140
tcatcatgtt ggccagggtg gtttcgaact cctgacctga ggtgagctgc ccacctggc 1200
ctcccaaagt gctgggatta caggcatgag ccactgccc cagctcagat ccatcccttt 1260
ctaagggcaa acagtccatg gtgcaaaggg gccatgccac ccagagttat gactacctgg 1320
gactccagaa ttccttgctt ggtggcctcc acatgcactt ccagggcctg cttgggcctc 1380
```

```
ttctatgcgt ctgtcctgag tgttgataga accactgatg tgagtacctg ggcttgagcc 1440
gtggcctgga gatcctgttg actgtagcat ggagggggct tgtgcagctg aatgtctgca 1500
tgcaggtggt gggagttctg gaatatgatg gagctggagg tgggaagaga agtaggcttg 1560
gggcagctct ctcattgccac ctcatctctg ccaaaactca ggtcaaactg tgaagagtct 1620
aaatgtgaat ctgcccttca aggtggctac aaaggtatct ttgtcaaggt aggagacctt 1680
gtggcctcca cgtgcacttc cagggcctgc ttgggcctct tctacgggtc tgtcctgagt 1740
cttctatgaa tccttcaggg cagattcata tttagactct tcacagtttg acctgagttt 1800
tggccagaat aaggtgacat ttagtttgtt ggcttgatgg atgacttaaa tatttagaca 1860
tggtgtgtag gcctgcattc ctactcttgc cttttttttt gcccctccag tgttttgggt 1920
agttttgctc ccctacagcc aaaggcaaac agagaagttg gaggtctgga gtggctacat 1980
aattttacac gactgcaatt ctctggctgc acttcacaaa tgtatacaaa ctaaatacaa 2040
gtcctgtgtt tttatcacag ggaggctgat caatataatg aaattaaaag ggggctggctc 2100
catattgttc tgtgtttttg tttgtttgtt ttgtttgttt ctttttttgt ttttgtggcc 2160
tccttcctct caatttatga agagaagcag taagatgttc ctctcgggtc ctctgaggga 2220
cctggggagc tcaggctggg aatctccaag gcagtaggtc gcctatcaaa aatcaaagtc 2280
caggtttgtg gggggaaaac aaaagcagcc cattaccagc aggactgtcc gccttccctc 2340
caccacagcc taggcctttg aaaggaaaca aaagacaaga caaatgatt ggcgtcctga 2400
gggagattca gcctagagct ctctctcccc caatccctcc ctccggctga ggaaactaac 2460
aaaggaaaaa aaaattgcgg aaagcaggat ttagaggaag caaattccac tgggtgccctt 2520
ggctgccggg aacgtggact agagagtctg cggcgcagcc ccgagcccag cgcttcccgc 2580
gcgtcttagg ccggcggggc cgggcggggg aaggggagcg agaccgcgga ccctaagaca 2640
cctgctgtac cctccacccc caccacccc caccacccc cccccaactc cctagatgtg 2700
tcgtgggcgg ctgaacgtcg cccgtttaag gggcggggcc cggctccacg tgctttctgc 2760
tgagtgactg aactacataa acagaggccg ggaagggggc ggggaggagg gagagcacag 2820
gctttgaccg atagtaacct ctgcgctcgg tgcagccgaa tctataaaag gaactagtcc 2880
cggcaaaaac ccc 2893
```

<210> 4

<211> 221

<212> DNA

<213> Homo sapiens

<400> 4

```
gtaattgcga gcgagagtga gtggggccgg gaccgcgaga gccgagccga cccttctctc 60
ccgggctgcg gcagggcagg gcggggagct ccgcgcacca acagagccgg ttctcagggc 120
gctttgctcc ttgttttttc cccggttctg ttttctcccc ttctccggaa ggcttgtcaa 180
ggggtaggag aaagagacgc aaacacaaaa gtggaaaaca g 221
```

<210> 5

<211> 159

<212> DNA

<213> Homo sapiens

<400> 5

```
ttaatgacca gccacgggcg tccctgctgt cagctctggc cgctgccttc cagggctccc 60
gagccacacg ctgggcgtgc tggctgaggg aacatggcat gttggcctca gctgaggttg 120
ctgctgtgga agaacctcac tttcagaaga agacaaaaca 159
```

<210> 6

<211> 117

<212> DNA

<213> Homo sapiens

<400> 6

```
gtaagaggct ctccagtgc ttaacttggc gttattgttt tgtttcgagg ccaaggaggc 60
ttcgggaagt gctcggtttc ggggaacttg atccggagcc ccacatcccc accactt 117
```

<210> 7
<211> 99
<212> DNA
<213> Homo sapiens

<400> 7
tggaggtctc agctgagagg gctggattag cagtcctcat tgggtgtatgg ctttgcagca 60
ataactgatg gctgtttccc ctctgcttt atctttcag 99

<210> 8
<211> 99
<212> DNA
<213> Homo sapiens

<400> 8
gtaagcttgg gtttttcagc agcgggggggt tctctcattt tttctttgtg gttttgagtt 60
ggggattgga ggaggaggagg agggaaggaa gctgtgttg 99

<210> 9
<211> 22
<212> PRT
<213> Homo sapiens

<400> 9
Met Ala Cys Trp Pro Gln Leu Arg Leu Leu Leu Trp Lys Asn Leu Thr
1 5 10 15

Phe Arg Arg Arg Gln Thr
20

<210> 10
<211> 9741
<212> DNA
<213> Homo sapiens

<400> 10
cttgtttttt ccccggttct gttttctccc cttctccgga aggcttgtca aggggtagga 60
gaaagagacg caaacacaaa agtggaaaac agttaatgac cagccacggc gtccctgctg 120
tgagctctgg ccgctgcctt ccagggtccc cgagccacac gctgggggtg ctggctgagg 180
gaacatggct tggtggcctc agctgagggt gctgctgtgg aagaacctca ctttcagaag 240
aagacaaaaca tgtcagctgc tgctggaagt ggcttgccct ctatttatct tctgatcct 300
gatctctggt cggctgagct acccacccta tgaacaacat gaatgccatt ttccaaataa 360
agccatgccc tctgcaggaa cacttccttg gggtcagggg attatctgta atgccaacaa 420
cccctgtttc cgttaccgga ctccctggga ggctcccgga gttgttgga aactttaacaa 480
atccattgtg gctgcctgt tctcagatgc tcggaggctt cttttataca gccagaaaga 540
caccagcatg aaggacatgc gcaaagtctt gagaacatta cagcagatca agaaatccag 600
ctcaaacttg aagcttcaag atttccctgg ggacaatgaa accttctctg ggttcctgta 660
tcacaacctc tctctcccaa agtctactgt ggacaagatg ctgagggtg atgtcattct 720
ccacaaggta tttttgcaag gctaccagtt acatttgaca agtctgtgca atggatcaaa 780
atcagaagag atgattcaac ttggtgacca agaagtttct gagctttgtg gcctaccaag 840
ggagaaactg gctgcagcag agcgagtact tcgttccaac atggacatcc tgaagccaat 900
cctgagaaca ctaaactcta catctccctt cccgagcaag gagctggccg aagccacaaa 960
aacattgctg catagtcttg ggactctggc ccaggagctg ttcagcatga gaagctggag 1020
tgacatgcga caggaggtga tgtttctgac caatgtgaac agctccagct cctccacca 1080
aatctaccag gctgtgtctc gtattgtctg cgggcatccc gagggagggg ggctgaagat 1140
caagtctctc aactggtatg aggacaacaa ctacaaagcc ctctttggag gcaatggcac 1200
tgaggaagat gctgaaacct tctatgacaa ctctacaact ccttactgca atgatttgat 1260

gaagaatttg	gagtctagtc	ctctttcccg	cattatctgg	aaagctctga	agccgctgct	1320
cgttgggaag	atcctgtata	cacctgacac	tccagccaca	aggcagggtca	tggtctgaggt	1380
gaacaagacc	ttccagggaac	tggtctgtgt	ccatgatctg	gaaggcatgt	gggaggaact	1440
cagccccaag	atctggacct	tcattggagaa	cagccaagaa	atggaccttg	tccggatgct	1500
gttggacagc	agggacaatg	accacttttg	ggaacagcag	ttggatggct	tagattggac	1560
agcccaagac	atcgtggcgt	ttttggccaa	gcaccagag	gatgtccagt	ccagtaatgg	1620
ttctgtgtac	acctggagag	aagctttcaa	cgagactaac	caggcaatcc	ggaccatata	1680
tcgcttcatg	gagtgtgtca	acctgaacaa	gctagaaccc	atagcaacag	aagtctggct	1740
catcaacaag	tccatggagc	tgctggatga	gaggaagttc	tgggctggta	ttgtgttcac	1800
tgggaattact	ccaggcagca	ttgagctgcc	ccatcatgtc	aagtacaaga	tccgaatgga	1860
cattgacaat	gtggagagga	caaataaaat	caaggatggg	tactgggacc	ctggtcctcg	1920
agctgacccc	tttgaggaca	tgcggtacgt	ctgggggggc	ttcgccctact	tgcaggatgt	1980
gggtggagcag	gcaatcatca	gggtgctgac	gggcaccgag	aagaaaactg	gtgtctatat	2040
gcaacagatg	ccctatccct	gttacgttga	tgacatcttt	ctgcgggtga	tgagccggct	2100
aatgcccctc	ttcatgacgc	tggcctggat	ttactcagtg	gctgtgatca	tcaagggcat	2160
cgtgtatgag	aaggaggcac	ggctgaaaga	gaccatgcgg	atcatgggccc	tggacaacag	2220
catcctcttg	tttagctggg	tcattagtag	cctcattcct	cttcttgtga	gcgctggcct	2280
gctagtggtc	atcctgaagt	taggaaacct	gctgccctac	agtgatccca	gcgtgggtgt	2340
tgtcttccctg	tccgtgtttg	ctgtgggtgac	aatcctgcag	tgcttccctga	ttagcacact	2400
cttctccaga	gccaaacctg	cagcagcctg	tggggggcatc	atctacttca	cgctgtacct	2460
gccctacgtc	ctgtgtgtgg	catggcagga	ctacgtgggc	ttcacactca	agatcttcgc	2520
tagcctgctg	tctcctgtgg	cttttggggt	tggtctgtgag	tactttgccc	tttttgagga	2580
gcagggcatt	ggagtgacgt	gggacaacct	gtttgagagt	cctgtggagg	aagatggcct	2640
caatctcacc	acttcgggtc	ccatgatgct	gtttgacacc	ttcctctatg	gggtgatgac	2700
ctggtacatt	gaggctgtct	ttccaggcca	gtacggaatt	cccaggccct	ggtattttcc	2760
ttgcaccaag	tctactggg	ttggcgagga	aagtgatgag	aagagccacc	ctggttccaa	2820
ccagaagaga	atatcagaaa	tctgcatgga	ggaggaaccc	accacttga	agctgggcgt	2880
gtccattcag	aacctggtaa	aagtctaccg	agatgggatg	aagggtggctg	tcgatggcct	2940
ggcactgaat	ttttatgagg	gccagatcac	ctccttccctg	ggccacaatg	gagcggggaa	3000
gacgaccacc	atgtcaatcc	tgaccgggtt	gttccccccg	acctcgggca	ccgcctatgc	3060
cctgggaaaa	gacattcgct	ctgagatgag	caccatccgg	cagaacctgg	gggtctgtcc	3120
ccagcataac	gtgctgtttg	acatgctgac	tgtcgaagaa	cacatctggt	tctatgcccg	3180
cttgaaaggg	ctctctgaga	agcacgtgaa	ggcggagatg	gagcagatgg	ccctggatgt	3240
tggtttgcca	tcaagcaagc	tgaaaagcaa	aacaagccag	ctgtcagggtg	gaatgcagag	3300
aaagctatct	gtggccttgg	cctttgtcgg	gggatctaag	gttgtcatte	tggatgaacc	3360
cacagctggg	gtggaccctt	actcccgcag	gggaatatgg	gagctgctgc	tgaataaccg	3420
acaaggccgc	accattatct	tctctacaca	ccacatggat	gaagcggacg	tcctggggga	3480
caggattgccc	atcatctccc	atgggaagct	gtgctgtgtg	ggctcctccc	tggttctgaa	3540
gaaccagctg	ggaacaggct	actacctgac	cttggtcaag	aaagatgtgg	aatcctccct	3600
cagttcctgc	agaaacagta	gtagcactgt	gtcatacctg	aaaaaggagg	acagtgtttc	3660
tcagagcagt	tctgatgctg	gcctgggcag	cgaccatgag	agtgacacgc	tgaccatcga	3720
tgtctctgct	atctccaacc	tcacagga	gcattgtgtc	gaagcccggc	tgggtggaaga	3780
catagggcat	gagctgacct	atgtgctgcc	atatgaagct	gctaaggagg	gagcctttgt	3840
ggaactcttt	catgagattg	atgaccggct	ctcagacctg	ggcattttcta	gttatggcat	3900
ctcagagacg	accctggaag	aaatattcct	caagggtggc	gaagagagtg	gggtggatgc	3960
tgagacctca	gatggtacct	tgccagcaag	acgaaacagg	cgggccttcg	gggacaagca	4020
cagctgtctt	cgcccgttca	ctgaagatga	tgtctgtgat	ccaaatgatt	ctgacataga	4080
cccagatacc	agagagacag	acttgctcag	tgggatggat	ggcaaagggt	cctaccagggt	4140
gaaaggctgg	aaacttacac	agcaacagtt	tgtggccctt	ttgtggaaga	gactgctaata	4200
tgccagacgg	agtcggaaag	gattttttgc	tcagattgtc	ttgccagctg	tggttgtctg	4260
cattgccctt	gtgttcagcc	tgatcgtgcc	accctttggc	aagtacccca	gcctggaact	4320
tcagccctgg	atgtacaacg	aacagtacac	atttgtcagc	aatgatgctc	ctgaggacac	4380
gggaaccctg	gaactcttaa	acgccctcac	caaagaccct	ggcttcggga	cccgtgtgat	4440
ggaaggaaac	ccaatcccag	acacgccctg	ccaggcaggg	gaggaagagt	ggaccactgc	4500
cccagttccc	cagaccatca	tggacctctt	ccagaatggg	aactggacaa	tgcagaacct	4560
ttcacctgca	tgccagtgtg	gcagcgacaa	aatcaagaag	atgtgcctg	tgtgtcccc	4620
aggggcaggg	gggctgcctc	ctccacaaac	aaaacaaaac	actgcagata	tccttcagga	4680
cctgacagga	agaaaacatt	cggattatct	gggtgaagacg	tatgtgcaga	tcatagccaa	4740
aagcttaaag	aacaagatct	gggtgaatga	gtttaggtat	ggcggctttt	ccctgggtgt	4800
cagtaatact	caagcacttc	ctccgagtca	agaagttaat	gatgccacca	aacaaatgaa	4860
gaaacaccta	aagctggcca	aggacagttc	tgcagatcga	tttctcaaca	gcttgggaag	4920

atttatgaca	ggactggaca	ccagaaataa	tgtcaagggtg	tggttcaata	acaagggctg	4980
gcatgcaatc	agctctttcc	tgaatgtcat	caacaatgcc	attctccggg	ccaacctgca	5040
aaagggagag	aaccctagcc	attatggaat	tactgctttc	aatcatcccc	tgaatctcac	5100
caagcagcag	ctctcagag	tggctccgat	gaccacatca	gtggatgtcc	ttgtgtccat	5160
ctgtgtcatc	tttgcaatgt	ccttcgtccc	agccagcttt	gtcgtattcc	tgatccagga	5220
gcggttcagc	aaagcaaaac	acctgcagtt	catcagtgga	gtgaagcctg	tcattctactg	5280
gctctcta	tttgtctggg	atatgtgcaa	ttacgttgct	cctgccacac	tggtcattat	5340
catcttcac	tgcttccagc	agaagtccta	tgtgtcctcc	accaatctgc	ctgtgctagc	5400
ccttctactt	ttgctgtatg	ggtggtcaat	cacacctctc	atgtaccag	cctcctttgt	5460
gttcaagatc	cccagcacag	cctatgtggt	gctcaccagc	gtgaacctct	tcattggcat	5520
taatggcagc	gtggccacct	ttgtgctgga	gctgttcacc	gacaataagc	tgaataatat	5580
caatgatata	ctgaagtcgg	tggtcttgat	cttcccacat	ttttgcctgg	gacgagggct	5640
catcgacatg	gtgaaaaacc	aggcaatggc	tgatgccctg	gaaagggttg	gggagaatcg	5700
ccttgtgtca	ccattatctt	gggacttggt	gggacgaaac	ctcttcgcca	tggccgtgga	5760
aggggtggtg	ttcttctca	ttactgttct	gatccagtac	agattcttca	tcaggcccag	5820
acctgtaaat	gcaagctat	ctcctctgaa	tgatgaagat	gaagatgtga	ggcgggaaag	5880
acagagaatt	cttgatggtg	gaggccagaa	tgacatctta	gaaatcaagg	agttgacgaa	5940
gatatataga	aggaagcggg	agcctgctgt	tgacaggatt	tgcgtgggca	ttcctcctgg	6000
tgagtgtctt	gggctcctgg	gagttaatgg	ggctggaaaa	tcatacaact	tcaagatgtt	6060
aacaggagat	accactgtta	ccagaggaga	tgctttcctt	aacagaaata	gtatcttatc	6120
aaacatccat	gaagtacatc	agaacatggg	ctactgccct	cagtttgatg	ccatcacaga	6180
gctgttgact	gggagataac	acgtggagtt	ctttgcctt	ttgagaggag	tcccagagaa	6240
agaagtgggc	aagggttggtg	agtggcgat	tgggaaactg	ggcctcgtga	agtatggaga	6300
aaaatatgct	ggtaactata	gtggaggcaa	caaacgcaag	ctctctacag	ccatggcttt	6360
gateggcggg	cctcctgtgg	tggttctgga	tgaaccacc	acaggcatgg	atcccaaagc	6420
ccggcggttc	ttgtggaatt	gtgccctaag	tggtgtcaag	gaggggagat	cagtagtgct	6480
tacatctcat	agtatggaag	aatgtgaagc	tctttgcact	aggatggcaa	tcatggtcaa	6540
tgggaaggtt	aggtgccttg	gcagtgtcca	gcacttaaaa	aatagggttg	gagatggtta	6600
tacaatagtt	gtacgaatag	cagggtccaa	cccggacctg	aagcctgtcc	aggatttctt	6660
tggacttgca	tttcttgga	gtgttccaaa	agagaaacac	cggaaacatgc	tacaatacca	6720
gcttccatct	tccttatctt	ctctggccag	gatattcagc	atcctctccc	agagcaaaaa	6780
gcgactccac	atagaagact	actctgtttc	tcagacaaca	cttgaccaag	tatttgtgaa	6840
ccttgccaag	gaccaaaagt	atgatgacca	cttaaaagac	ctctcattac	acaaaaacca	6900
gacagtagtg	gacgttgacg	ttctcacatc	ttttctacag	gatgagaaag	tgaaagaaag	6960
ctatgtatga	agaatcctgt	tcatacgggg	tggctgaaag	taaagaggna	ctagactttc	7020
ccttgccacca	tgtgaagtgt	tgtggagaaa	agagccagaa	gttgatgtgg	gaagaagtaa	7080
actggatact	gtactgatac	tattcaatgc	aatgcaattc	aatgcaatga	aaacaaaatt	7140
ccattacagg	ggcagtgctt	ttgtagccta	tgtcttgat	ggctctcaag	tgaaagactt	7200
gaatttagtt	ttttacctat	acctatgtga	aactctatta	tggaaaccaa	tggacatatg	7260
ggtttgaact	cacacttttt	tttttttttt	gttctgtgt	attctcattg	gggttgcaac	7320
aataattcat	caagtaatca	tggccagcga	ttattgatca	aaatcaaaag	gtaatgcaca	7380
tcctcattca	ctaagccatg	ccatgcccag	gagactgggt	tcccggtgac	acatccattg	7440
ctggcaatga	gtgtgccaga	gttattagtg	ccaagttttt	cagaaagttt	gaagcaccat	7500
ggtgtgtcat	gtcacttttt	gtgaaagctg	ctctgctcag	agtctatcaa	cattgaatat	7560
cagttgacag	aatggtgcca	tgcgtggcta	acatcctgct	ttgattccct	ctgataagct	7620
gttctgggtg	cagtaacatg	caacaaaaat	gtgggtgtct	ctaggcacgg	gaaacttggt	7680
tccattgtta	tattgtccta	tgcttcgagc	catgggtcta	cagggtcata	ccttatgagac	7740
tcttaaatat	acttagatcc	tggtaagagg	caaagaatca	acagccaaac	tgctggggct	7800
gcaagctgct	gaagccaggg	catgggatta	aagagattgt	gcgttcaaac	ctaggggaagc	7860
ctgtgcccac	ttgtcctgac	tgtctgctaa	catggtacac	tgcatctcaa	gatgtttatc	7920
tgacacaagt	gtattatttc	tggctttttg	aattaatcta	gaaaatgaaa	agatggagtt	7980
gtattttgac	aaaaatgttt	gtacttttta	atgttatttg	gaattttaag	ttctatcagt	8040
gacttctgaa	tccttagaat	ggcctctttg	tagaaccctg	tggatatagag	gagtatggcc	8100
actgccccac	tattttttat	ttcttatgta	agtttgcata	tcagtcatga	ctagtgccta	8160
gaaagcaatg	tgatggctcag	gatctcatga	cattatattt	gagtttcttt	cagatcattt	8220
aggatactct	taatctcact	tcatacatca	aattattttt	gagtgtatgc	tgtagctgaa	8280
agagtatgta	cgtacgtata	agactagaga	gatattaagt	ctcagtacac	ttcctgtgcc	8340
atgttattca	gctcactggg	ttacaaatat	agggtgtctt	gtggttgtag	gagcccactg	8400
taacaatact	gggcagcctt	tttttttttt	tttaattgca	acaatgcaaa	agccaagaaa	8460
gtataagggt	cacaagtcta	aacaatgaat	tcttcaacag	ggaaaacagc	tagcttgaaa	8520
acttgctgaa	aaacacaact	tgtgtttatg	gcatttagta	ccttcaaata	attggctttg	8580

```

cagatattgg atacccatt aaatctgaca gtctcaaatt tttcatctct tcaatcacta 8640
gtcaagaaaa atataaaaaac aacaaaatact tccatatgga gcattttttca gagtttttcta 8700
acccagtctt attttttctag tcagtaaaca tttgtaaaaa tactgtttca ctaataactta 8760
ctgttaactg tcttgagaga aaagaaaaat atgagagAAC tattgtttgg ggaagttcaa 8820
gtgatctttc aatatcatta ctaacttctt ccactttttc caaaatttga atattaacgc 8880
taaaggtgta agacttcaga tttcaaatta atctttctat atttttttaa tttacagaat 8940
attatataac ccactgctga aaaagaaaaa aatgattgtt ttagaagtta aagtcaatat 9000
tgattttttaa tataagtaat gaaggcatat ttccaataac tagtgatatg gcacggttgc 9060
attttacagt atcttcaaaa atacagaatt tatagaataa tttctcctca tttaatattt 9120
ttcaaaatca aagttatggg ttcttcattt tactaaaatc gtatttcta tcttcattat 9180
agtaaactta tgagcaactc cttacttcgg ttctctgat ttcaaggcca tatttttaaaa 9240
aatcaaaagg cactgtgaac tattttgaag aaaacacaac attttaatac agattgaaag 9300
gacctcttct gaagctagaa acaatctata gttatacatc ttcattaata ctgtgttacc 9360
ttttaaaata gtaatttttt acatttttct gtgtaaacct aattgtggta gaaattttta 9420
ccaactctat actcaatcaa gcaaaatttc tgtatatcc ctgtggaatg tacctatgtg 9480
agtttcagaa attctcaaaa tacgtgttca aaaatttctg cttttgcatc tttgggacac 9540
ctcagaaaac ttattaacaa ctgtgaatat gagaaatac gaagaaaata ataagccctc 9600
tatacataaa tgcccagcac aattcattgt taaaaaaca ccaaacctca cactactgta 9660
tttcattatc tgtactgaaa gcaaatgctt tgtgactatt aaatgttgca catcattcat 9720
tcaaaaaaaaa aaaaaaaaaa a 9741

```

<210> 11

<211> 2261

<212> PRT

<213> Homo sapiens

<400> 11

```

Met Ala Cys Trp Pro Gln Leu Arg Leu Leu Leu Trp Lys Asn Leu Thr
  1              5              10              15

Phe Arg Arg Arg Gln Thr Cys Gln Leu Leu Leu Glu Val Ala Trp Pro
      20              25              30

Leu Phe Ile Phe Leu Ile Leu Ile Ser Val Arg Leu Ser Tyr Pro Pro
      35              40              45

Tyr Glu Gln His Glu Cys His Phe Pro Asn Lys Ala Met Pro Ser Ala
      50              55              60

Gly Thr Leu Pro Trp Val Gln Gly Ile Ile Cys Asn Ala Asn Asn Pro
      65              70              75              80

Cys Phe Arg Tyr Pro Thr Pro Gly Glu Ala Pro Gly Val Val Gly Asn
      85              90              95

Phe Asn Lys Ser Ile Val Ala Arg Leu Phe Ser Asp Ala Arg Arg Leu
      100             105             110

Leu Leu Tyr Ser Gln Lys Asp Thr Ser Met Lys Asp Met Arg Lys Val
      115             120             125

Leu Arg Thr Leu Gln Gln Ile Lys Lys Ser Ser Ser Asn Leu Lys Leu
      130             135             140

Gln Asp Phe Leu Val Asp Asn Glu Thr Phe Ser Gly Phe Leu Tyr His
      145             150             155             160

Asn Leu Ser Leu Pro Lys Ser Thr Val Asp Lys Met Leu Arg Ala Asp
      165             170             175

```


Val Ile Leu His Lys Val Phe Leu Gln Gly Tyr Gln Leu His Leu Thr
 180 185 190
 Ser Leu Cys Asn Gly Ser Lys Ser Glu Glu Met Ile Gln Leu Gly Asp
 195 200 205
 Gln Glu Val Ser Glu Leu Cys Gly Leu Pro Arg Glu Lys Leu Ala Ala
 210 215 220
 Ala Glu Arg Val Leu Arg Ser Asn Met Asp Ile Leu Lys Pro Ile Leu
 225 230 235 240
 Arg Thr Leu Asn Ser Thr Ser Pro Phe Pro Ser Lys Glu Leu Ala Glu
 245 250 255
 Ala Thr Lys Thr Leu Leu His Ser Leu Gly Thr Leu Ala Gln Glu Leu
 260 265 270
 Phe Ser Met Arg Ser Trp Ser Asp Met Arg Gln Glu Val Met Phe Leu
 275 280 285
 Thr Asn Val Asn Ser Ser Ser Ser Ser Thr Gln Ile Tyr Gln Ala Val
 290 295 300
 Ser Arg Ile Val Cys Gly His Pro Glu Gly Gly Gly Leu Lys Ile Lys
 305 310 315 320
 Ser Leu Asn Trp Tyr Glu Asp Asn Asn Tyr Lys Ala Leu Phe Gly Gly
 325 330 335
 Asn Gly Thr Glu Glu Asp Ala Glu Thr Phe Tyr Asp Asn Ser Thr Thr
 340 345 350
 Pro Tyr Cys Asn Asp Leu Met Lys Asn Leu Glu Ser Ser Pro Leu Ser
 355 360 365
 Arg Ile Ile Trp Lys Ala Leu Lys Pro Leu Leu Val Gly Lys Ile Leu
 370 375 380
 Tyr Thr Pro Asp Thr Pro Ala Thr Arg Gln Val Met Ala Glu Val Asn
 385 390 395 400
 Lys Thr Phe Gln Glu Leu Ala Val Phe His Asp Leu Glu Gly Met Trp
 405 410 415
 Glu Glu Leu Ser Pro Lys Ile Trp Thr Phe Met Glu Asn Ser Gln Glu
 420 425 430
 Met Asp Leu Val Arg Met Leu Leu Asp Ser Arg Asp Asn Asp His Phe
 435 440 445
 Trp Glu Gln Gln Leu Asp Gly Leu Asp Trp Thr Ala Gln Asp Ile Val
 450 455 460
 Ala Phe Leu Ala Lys His Pro Glu Asp Val Gln Ser Ser Asn Gly Ser
 465 470 475 480
 Val Tyr Thr Trp Arg Glu Ala Phe Asn Glu Thr Asn Gln Ala Ile Arg
 485 490 495
 Thr Ile Ser Arg Phe Met Glu Cys Val Asn Leu Asn Lys Leu Glu Pro

500	505	510
Ile Ala Thr Glu Val Trp Leu Ile Asn Lys Ser Met Glu Leu Leu Asp 515	520	525
Glu Arg Lys Phe Trp Ala Gly Ile Val Phe Thr Gly Ile Thr Pro Gly 530	535	540
Ser Ile Glu Leu Pro His His Val Lys Tyr Lys Ile Arg Met Asp Ile 545	550	555
Asp Asn Val Glu Arg Thr Asn Lys Ile Lys Asp Gly Tyr Trp Asp Pro 565	570	575
Gly Pro Arg Ala Asp Pro Phe Glu Asp Met Arg Tyr Val Trp Gly Gly 580	585	590
Phe Ala Tyr Leu Gln Asp Val Val Glu Gln Ala Ile Ile Arg Val Leu 595	600	605
Thr Gly Thr Glu Lys Lys Thr Gly Val Tyr Met Gln Gln Met Pro Tyr 610	615	620
Pro Cys Tyr Val Asp Asp Ile Phe Leu Arg Val Met Ser Arg Ser Met 625	630	635
Pro Leu Phe Met Thr Leu Ala Trp Ile Tyr Ser Val Ala Val Ile Ile 645	650	655
Lys Gly Ile Val Tyr Glu Lys Glu Ala Arg Leu Lys Glu Thr Met Arg 660	665	670
Ile Met Gly Leu Asp Asn Ser Ile Leu Trp Phe Ser Trp Phe Ile Ser 675	680	685
Ser Leu Ile Pro Leu Leu Val Ser Ala Gly Leu Leu Val Val Ile Leu 690	695	700
Lys Leu Gly Asn Leu Leu Pro Tyr Ser Asp Pro Ser Val Val Phe Val 705	710	715
Phe Leu Ser Val Phe Ala Val Val Thr Ile Leu Gln Cys Phe Leu Ile 725	730	735
Ser Thr Leu Phe Ser Arg Ala Asn Leu Ala Ala Ala Cys Gly Gly Ile 740	745	750
Ile Tyr Phe Thr Leu Tyr Leu Pro Tyr Val Leu Cys Val Ala Trp Gln 755	760	765
Asp Tyr Val Gly Phe Thr Leu Lys Ile Phe Ala Ser Leu Leu Ser Pro 770	775	780
Val Ala Phe Gly Phe Gly Cys Glu Tyr Phe Ala Leu Phe Glu Glu Gln 785	790	795
Gly Ile Gly Val Gln Trp Asp Asn Leu Phe Glu Ser Pro Val Glu Glu 805	810	815
Asp Gly Phe Asn Leu Thr Thr Ser Val Ser Met Met Leu Phe Asp Thr 820	825	830

Phe Leu Tyr Gly Val Met Thr Trp Tyr Ile Glu Ala Val Phe Pro Gly
 835 840 845
 Gln Tyr Gly Ile Pro Arg Pro Trp Tyr Phe Pro Cys Thr Lys Ser Tyr
 850 855 860
 Trp Phe Gly Glu Glu Ser Asp Glu Lys Ser His Pro Gly Ser Asn Gln
 865 870 875 880
 Lys Arg Ile Ser Glu Ile Cys Met Glu Glu Glu Pro Thr His Leu Lys
 885 890 895
 Leu Gly Val Ser Ile Gln Asn Leu Val Lys Val Tyr Arg Asp Gly Met
 900 905 910
 Lys Val Ala Val Asp Gly Leu Ala Leu Asn Phe Tyr Glu Gly Gln Ile
 915 920 925
 Thr Ser Phe Leu Gly His Asn Gly Ala Gly Lys Thr Thr Thr Met Ser
 930 935 940
 Ile Leu Thr Gly Leu Phe Pro Pro Thr Ser Gly Thr Ala Tyr Ile Leu
 945 950 955 960
 Gly Lys Asp Ile Arg Ser Glu Met Ser Thr Ile Arg Gln Asn Leu Gly
 965 970 975
 Val Cys Pro Gln His Asn Val Leu Phe Asp Met Leu Thr Val Glu Glu
 980 985 990
 His Ile Trp Phe Tyr Ala Arg Leu Lys Gly Leu Ser Glu Lys His Val
 995 1000 1005
 Lys Ala Glu Met Glu Gln Met Ala Leu Asp Val Gly Leu Pro Ser Ser
 1010 1015 1020
 Lys Leu Lys Ser Lys Thr Ser Gln Leu Ser Gly Gly Met Gln Arg Lys
 1025 1030 1035 1040
 Leu Ser Val Ala Leu Ala Phe Val Gly Gly Ser Lys Val Val Ile Leu
 1045 1050 1055
 Asp Glu Pro Thr Ala Gly Val Asp Pro Tyr Ser Arg Arg Gly Ile Trp
 1060 1065 1070
 Glu Leu Leu Leu Lys Tyr Arg Gln Gly Arg Thr Ile Ile Leu Ser Thr
 1075 1080 1085
 His His Met Asp Glu Ala Asp Val Leu Gly Asp Arg Ile Ala Ile Ile
 1090 1095 1100
 Ser His Gly Lys Leu Cys Cys Val Gly Ser Ser Leu Phe Leu Lys Asn
 1105 1110 1115 1120
 Gln Leu Gly Thr Gly Tyr Tyr Leu Thr Leu Val Lys Lys Asp Val Glu
 1125 1130 1135
 Ser Ser Leu Ser Ser Cys Arg Asn Ser Ser Ser Thr Val Ser Tyr Leu
 1140 1145 1150

Lys Lys Glu Asp Ser Val Ser Gln Ser Ser Ser Asp Ala Gly Leu Gly
1155 1160 1165

Ser Asp His Glu Ser Asp Thr Leu Thr Ile Asp Val Ser Ala Ile Ser
1170 1175 1180

Asn Leu Ile Arg Lys His Val Ser Glu Ala Arg Leu Val Glu Asp Ile
1185 1190 1195 1200

Gly His Glu Leu Thr Tyr Val Leu Pro Tyr Glu Ala Ala Lys Glu Gly
1205 1210 1215

Ala Phe Val Glu Leu Phe His Glu Ile Asp Asp Arg Leu Ser Asp Leu
1220 1225 1230

Gly Ile Ser Ser Tyr Gly Ile Ser Glu Thr Thr Leu Glu Glu Ile Phe
1235 1240 1245

Leu Lys Val Ala Glu Glu Ser Gly Val Asp Ala Glu Thr Ser Asp Gly
1250 1255 1260

Thr Leu Pro Ala Arg Arg Asn Arg Arg Ala Phe Gly Asp Lys Gln Ser
1265 1270 1275 1280

Cys Leu Arg Pro Phe Thr Glu Asp Asp Ala Ala Asp Pro Asn Asp Ser
1285 1290 1295

Asp Ile Asp Pro Glu Ser Arg Glu Thr Asp Leu Leu Ser Gly Met Asp
1300 1305 1310

Gly Lys Gly Ser Tyr Gln Val Lys Gly Trp Lys Leu Thr Gln Gln Gln
1315 1320 1325

Phe Val Ala Leu Leu Trp Lys Arg Leu Leu Ile Ala Arg Arg Ser Arg
1330 1335 1340

Lys Gly Phe Phe Ala Gln Ile Val Leu Pro Ala Val Phe Val Cys Ile
1345 1350 1355 1360

Ala Leu Val Phe Ser Leu Ile Val Pro Pro Phe Gly Lys Tyr Pro Ser
1365 1370 1375

Leu Glu Leu Gln Pro Trp Met Tyr Asn Glu Gln Tyr Thr Phe Val Ser
1380 1385 1390

Asn Asp Ala Pro Glu Asp Thr Gly Thr Leu Glu Leu Leu Asn Ala Leu
1395 1400 1405

Thr Lys Asp Pro Gly Phe Gly Thr Arg Cys Met Glu Gly Asn Pro Ile
1410 1415 1420

Pro Asp Thr Pro Cys Gln Ala Gly Glu Glu Glu Trp Thr Thr Ala Pro
1425 1430 1435 1440

Val Pro Gln Thr Ile Met Asp Leu Phe Gln Asn Gly Asn Trp Thr Met
1445 1450 1455

Gln Asn Pro Ser Pro Ala Cys Gln Cys Ser Ser Asp Lys Ile Lys Lys
1460 1465 1470

Met Leu Pro Val Cys Pro Pro Gly Ala Gly Gly Leu Pro Pro Pro Gln

1475	1480	1485
Arg Lys Gln Asn Thr Ala Asp Ile Leu Gln Asp Leu Thr Gly Arg Asn 1490	1495	1500
Ile Ser Asp Tyr Leu Val Lys Thr Tyr Val Gln Ile Ile Ala Lys Ser 1505	1510	1515 1520
Leu Lys Asn Lys Ile Trp Val Asn Glu Phe Arg Tyr Gly Gly Phe Ser 1525	1530	1535
Leu Gly Val Ser Asn Thr Gln Ala Leu Pro Pro Ser Gln Glu Val Asn 1540	1545	1550
Asp Ala Thr Lys Gln Met Lys Lys His Leu Lys Leu Ala Lys Asp Ser 1555	1560	1565
Ser Ala Asp Arg Phe Leu Asn Ser Leu Gly Arg Phe Met Thr Gly Leu 1570	1575	1580
Asp Thr Arg Asn Asn Val Lys Val Trp Phe Asn Asn Lys Gly Trp His 1585	1590	1595 1600
Ala Ile Ser Ser Phe Leu Asn Val Ile Asn Asn Ala Ile Leu Arg Ala 1605	1610	1615
Asn Leu Gln Lys Gly Glu Asn Pro Ser His Tyr Gly Ile Thr Ala Phe 1620	1625	1630
Asn His Pro Leu Asn Leu Thr Lys Gln Gln Leu Ser Glu Val Ala Pro 1635	1640	1645
Met Thr Thr Ser Val Asp Val Leu Val Ser Ile Cys Val Ile Phe Ala 1650	1655	1660
Met Ser Phe Val Pro Ala Ser Phe Val Val Phe Leu Ile Gln Glu Arg 1665	1670	1675 1680
Val Ser Lys Ala Lys His Leu Gln Phe Ile Ser Gly Val Lys Pro Val 1685	1690	1695
Ile Tyr Trp Leu Ser Asn Phe Val Trp Asp Met Cys Asn Tyr Val Val 1700	1705	1710
Pro Ala Thr Leu Val Ile Ile Ile Phe Ile Cys Phe Gln Gln Lys Ser 1715	1720	1725
Tyr Val Ser Ser Thr Asn Leu Pro Val Leu Ala Leu Leu Leu Leu 1730	1735	1740
Tyr Gly Trp Ser Ile Thr Pro Leu Met Tyr Pro Ala Ser Phe Val Phe 1745	1750	1755 1760
Lys Ile Pro Ser Thr Ala Tyr Val Val Leu Thr Ser Val Asn Leu Phe 1765	1770	1775
Ile Gly Ile Asn Gly Ser Val Ala Thr Phe Val Leu Glu Leu Phe Thr 1780	1785	1790
Asp Asn Lys Leu Asn Asn Ile Asn Asp Ile Leu Lys Ser Val Phe Leu 1795	1800	1805

Ile Phe Pro His Phe Cys Leu Gly Arg Gly Leu Ile Asp Met Val Lys
 1810 1815 1820
 Asn Gln Ala Met Ala Asp Ala Leu Glu Arg Phe Gly Glu Asn Arg Phe
 1825 1830 1835 1840
 Val Ser Pro Leu Ser Trp Asp Leu Val Gly Arg Asn Leu Phe Ala Met
 1845 1850 1855
 Ala Val Glu Gly Val Val Phe Phe Leu Ile Thr Val Leu Ile Gln Tyr
 1860 1865 1870
 Arg Phe Phe Ile Arg Pro Arg Pro Val Asn Ala Lys Leu Ser Pro Leu
 1875 1880 1885
 Asn Asp Glu Asp Glu Asp Val Arg Arg Glu Arg Gln Arg Ile Leu Asp
 1890 1895 1900
 Gly Gly Gly Gln Asn Asp Ile Leu Glu Ile Lys Glu Leu Thr Lys Ile
 1905 1910 1915 1920
 Tyr Arg Arg Lys Arg Lys Pro Ala Val Asp Arg Ile Cys Val Gly Ile
 1925 1930 1935
 Pro Pro Gly Glu Cys Phe Gly Leu Leu Gly Val Asn Gly Ala Gly Lys
 1940 1945 1950
 Ser Ser Thr Phe Lys Met Leu Thr Gly Asp Thr Thr Val Thr Arg Gly
 1955 1960 1965
 Asp Ala Phe Leu Asn Arg Asn Ser Ile Leu Ser Asn Ile His Glu Val
 1970 1975 1980
 His Gln Asn Met Gly Tyr Cys Pro Gln Phe Asp Ala Ile Thr Glu Leu
 1985 1990 1995 2000
 Leu Thr Gly Arg Glu His Val Glu Phe Phe Ala Leu Leu Arg Gly Val
 2005 2010 2015
 Pro Glu Lys Glu Val Gly Lys Val Gly Glu Trp Ala Ile Arg Lys Leu
 2020 2025 2030
 Gly Leu Val Lys Tyr Gly Glu Lys Tyr Ala Gly Asn Tyr Ser Gly Gly
 2035 2040 2045
 Asn Lys Arg Lys Leu Ser Thr Ala Met Ala Leu Ile Gly Gly Pro Pro
 2050 2055 2060
 Val Val Phe Leu Asp Glu Pro Thr Thr Gly Met Asp Pro Lys Ala Arg
 2065 2070 2075 2080
 Arg Phe Leu Trp Asn Cys Ala Leu Ser Val Val Lys Glu Gly Arg Ser
 2085 2090 2095
 Val Val Leu Thr Ser His Ser Met Glu Glu Cys Glu Ala Leu Cys Thr
 2100 2105 2110
 Arg Met Ala Ile Met Val Asn Gly Arg Phe Arg Cys Leu Gly Ser Val
 2115 2120 2125

Gln His Leu Lys Asn Arg Phe Gly Asp Gly Tyr Thr Ile Val Val Arg
 2130 2135 2140
 Ile Ala Gly Ser Asn Pro Asp Leu Lys Pro Val Gln Asp Phe Phe Gly
 2145 2150 2155 2160
 Leu Ala Phe Pro Gly Ser Val Pro Lys Glu Lys His Arg Asn Met Leu
 2165 2170 2175
 Gln Tyr Gln Leu Pro Ser Ser Leu Ser Ser Leu Ala Arg Ile Phe Ser
 2180 2185 2190
 Ile Leu Ser Gln Ser Lys Lys Arg Leu His Ile Glu Asp Tyr Ser Val
 2195 2200 2205
 Ser Gln Thr Thr Leu Asp Gln Val Phe Val Asn Phe Ala Lys Asp Gln
 2210 2215 2220
 Ser Asp Asp Asp His Leu Lys Asp Leu Ser Leu His Lys Asn Gln Thr
 2225 2230 2235 2240
 Val Val Asp Val Ala Val Leu Thr Ser Phe Leu Gln Asp Glu Lys Val
 2245 2250 2255
 Lys Glu Ser Tyr Val
 2260

<210> 12
 <211> 26
 <212> DNA
 <213> Homo sapiens

<400> 12
 ttgccgtcga ctgttttggg tagttt

26

<210> 13
 <211> 25
 <212> DNA
 <213> Homo sapiens

<400> 13
 gccctgtcga ccggtctgt tggtg

25

<210> 14
 <211> 29
 <212> DNA
 <213> Homo sapiens

<400> 14
 tcgcccgttt aggtttgggc gcccggtc

29

<210> 15
 <211> 29
 <212> DNA
 <213> Homo sapiens

<400> 15
cagaggccgg gaggccttggg cgggagggga

29

<210> 16
<211> 28
<212> DNA
<213> Homo sapiens

<400> 16
cgtgctttct gctgaggatg cgaactac

28

<210> 17
<211> 26
<212> DNA
<213> Homo sapiens

<400> 17
cggctcctca cggctttctg ctgagt

26

<210> 18
<211> 24
<212> DNA
<213> Homo sapiens

<400> 18
gcctcctttc tgctgagtga ctga

24

<210> 19
<211> 34
<212> DNA
<213> Homo sapiens

<400> 19
ctttgtgtga tagtaaacta ctgcgctcgg tgca

34

<210> 20
<211> 20
<212> DNA
<213> Homo sapiens

<400> 20
actcccaagc tttgtcgtgg

20

Figure 1

-1941
TTATTTTTCTTGACACGGAGTCTTGCTCTGTCACTCAGGCTGGAGTGCAGTGGCATGATC

-1881
TCGACTCACTGTAACCTCTGCCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCTGA

-1821 GATA
GATAACAGGCGCCCCGCCACCACATCTGGCTAATTTTTGTATTTTTAGTAAAGACTGGGTT

-1761
TCATCATGTTGGCCAGGTTGGTTTCGAACTCTGACCTGAGGTGAGCTGCCACCTTGGC
LXR α / Δ EF1 Δ EF1/

-1701 LYF1/IK2
CTCCCAAAGTGCTGGGATTACAGGCATGAGCCACTGCGCCCAGCTCAGATCCATCCCTTT

-1641
CTAAGGGCAAACAGTCCATGGTGCAAAGGGGCCATGCCACCCAGAGTTATGAGTACCTGG

-1581
GACTCCAGAATTCCTTGCCTGGTGGCCTCCACATGCACTTCCAGGGCCTGCTTGGGCCTC

-1521
TTCTATGCGTCTGTCCTGAGTGTTGATAGAACCACTGATGTGAGTACCTGGGCTTGAGCC

-1461
GTGGCCTGGAGATCCTGTTGACTGTAGCATGGAGGGGGCTTGTCAGCTGAATGTCTGCA
AP4 LMO2-

-1401 COM/MYOD/ Δ EF1/E47
TGCAGGTGGTGGGA GTTCTGGAATATGATGGAGCTGGAGGTGGGAAGAGAAGTAGGCTTG
ZID/ Δ EF1

-1341
GGGCAGCTCTCTCATGCCACCTCATTCTGGCCAAAACTCAGGTCAAACTGTGAAGAGTCT
 Δ EF1

-1281 PPAR PPAR
AAATGTGAATCTGCCCTTCAAGGTGGCTACAAAGGTATCTTTGTCAAGGTAGGAGACCTT

-1221 USF/NMYC/MYCMAX
GTGGCCTCCACGTGCACTTCCAGGGCCTGCTTGGGCCTCTTCTACGGGTCTGTCCTGAGT

-1161
CTTCTATGAATCCTTCAGGGCAGATTCATATTTAGACTCTTCACAGTTTGACCTG
 Δ EF1

-1101
AGTTTTGGCCAGAAATAAGGTGACATTAGTTTGTGGCTTGATCGATGACTTAAATATTT
 Δ EF1 SRY API

-1041
AGACATGGTGTGTAGGCCTGCATTCTACTCTTGCCTTTTTTTTTGGCCCTCCAGTGTTT

-981
TGGGTAGTTTTGCTCCCTACAGCCAAGGGCAAACAGAGAA GTTGGAGGTCTGGAGTGG
HNF3 β

Figure 1 (Suite 1)

-921 NKX2.5 PPAR/NKX2.5/PPAR
CTA CATAATTI TACACGACTGCAATTCTCTGGGTGCACTTCA CA AATGTATACAAACTAA

-861 GATA
ATACAAGTCCTGTGTI TTTATCACA GGGAGGCTGATCAATATAATGAAATTAAGGGGGG

-801 SOX5 SRY/HFH/HNF3 β SRY/HFH/HNF3 β
CTGGTCCAT ATTGTTCT GT TTTTGTITGTTTGTITGTTTGTITGTTTCTTTTTTTGTTT

-741
TGTGGCCTCCTTCCTCTCAATTTATGAAGAGAAGCAGTAAGATGTTCTCTCGGGTCCTC

-681 MZF1 IK2/NF κ B/CREL LMO2COM/GATA
TGAGGGA CCTGGGGA GCTCA GGCTGGGAATCTCCAA GGCAGTAGG TCGCCTATCAAAAAT

-621 MZF1/SRY PPAR PPAR
CAAAGTCCAGGTTTG TGGGGGGAAAACAAAAGCAGCCCA TTACCCAG AGGACT GTCCGCC

-561 MZF1 HNF3 β /SRY/EV11
T TCCCCTCA CCCCAGCCTAGGCCTTTG AAAGGAAACAAAAGACAAATGATTGGC

-501 MZF1 AP4
GTCCTGAGGGAGATTCAGCCTAGAGCTCTCTC TCCCCCA TCCCTCC TCCGGCTGAGGA

-441 SRY STAT
A ACTAACAAAGGA AAAAAAAATTGCGGAAAGCAGGAT TTAGAGGA AGCAAATTCCTG

-381 STAT/PPAR PPAR
TGCCCTTGCTG TCCG GGAACGT G GACTA GAGAGTCTGCGGCGCAGCCCCGAGCCCAGCGC

-321 AP2 MZF1
TTCCCGCGCTCTTA GCGCGCGGGCC CGGGCGGGG GAAGGGGA CGCAGACCGCGGACCC

-261 LMO2COM/MYOD/ E47 RREB1 MZF1/
TAAGACACCTGCTGT ACCCTCCA CCCCCACCCACCCCA CCCACCT CCCCCCAACTCCCT

-201 CMYB SPI/GC USF/NMYC/
AGATG TGCTGTGGGCGGCTGAACGTCGCCCCGTTT AAGGGGCGGGCCCG GCTCCACGTGC

-141 ARNT NFE2/AP1 XFD1/HFH GC/SPI/MZF1
TTTC TGCTGAGTGACTGAAC TACATAAACAGAGG CCGGGA AGGGGGCGGGGAG GAGGGAG

-81 LXR TATA
AGCACAGGCTTTG ACCGATAGTAACCTC TCGCTCGGTGCAGCCGAA TCTATAAAAG GAA

-21 +1
CTAGTCCCGGCAAAAACCCCG TAATTGCGAGCGAGAG

Figure 1 (suite 2)

FIGURE 2A

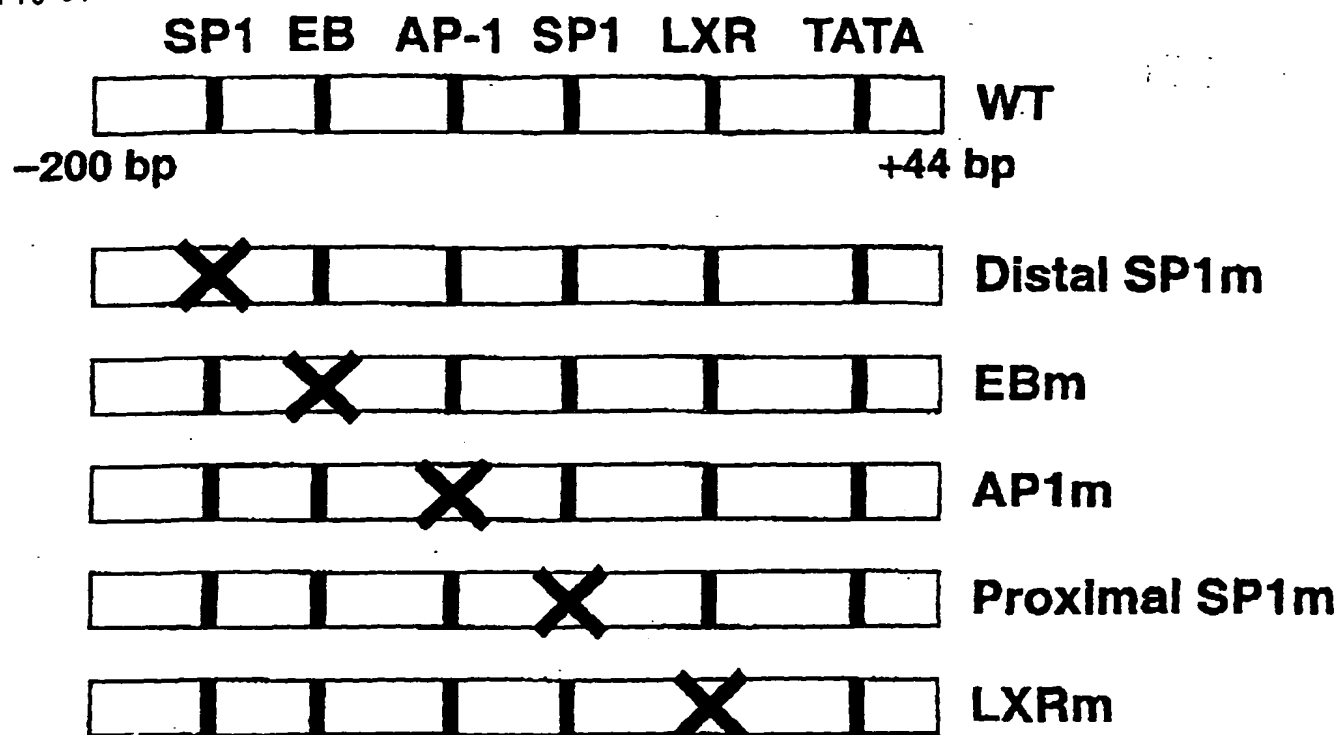


FIGURE 2B

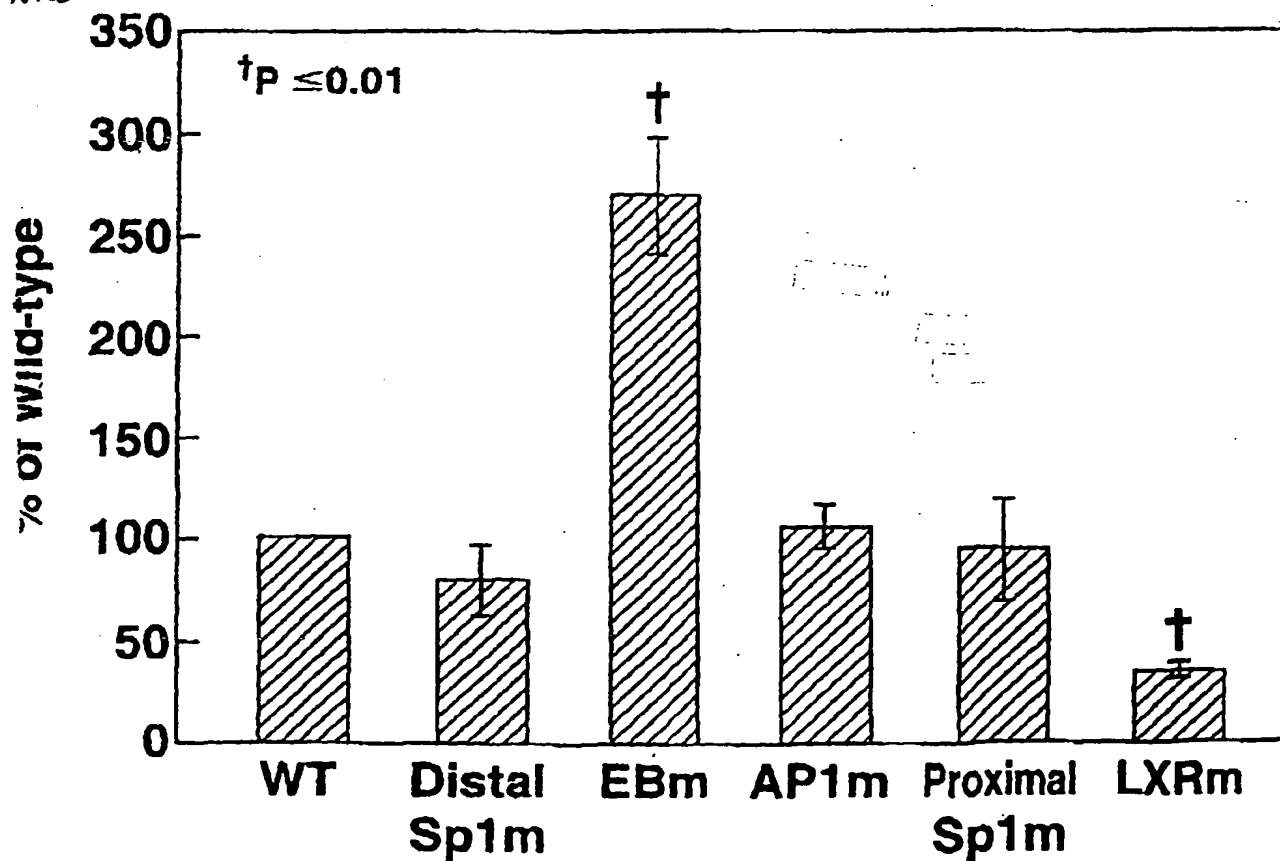
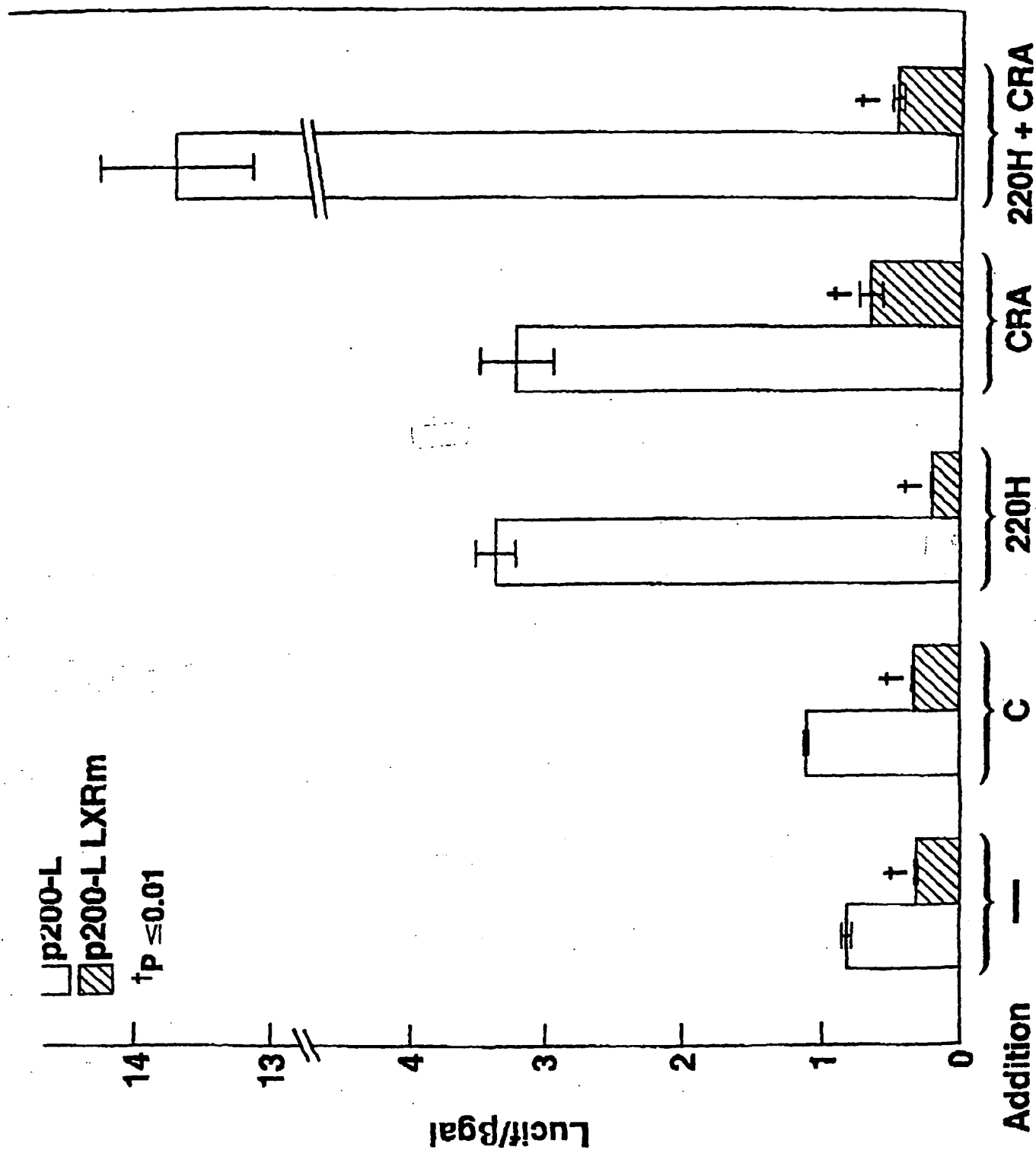


FIGURE 3



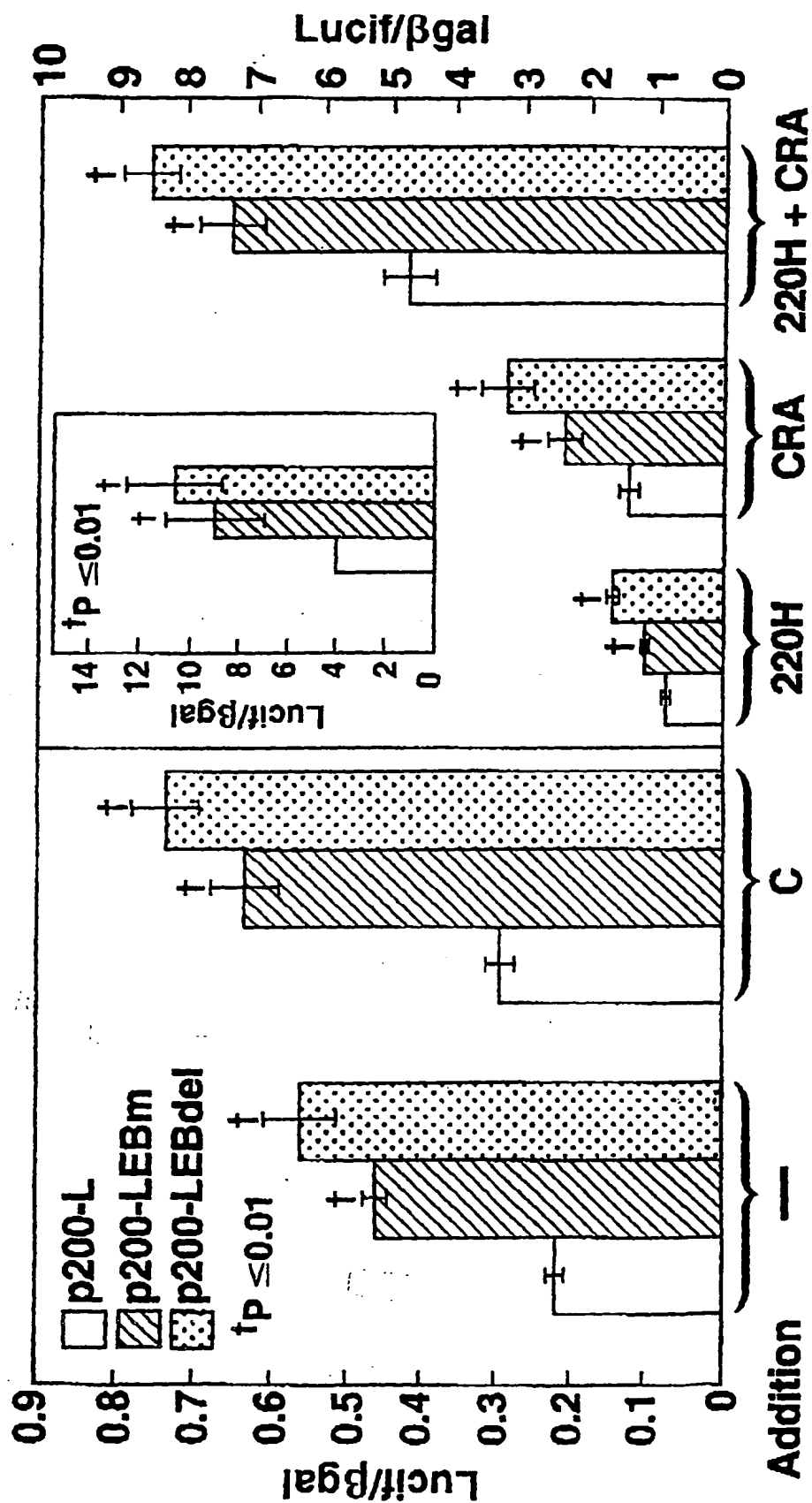


Fig. 4

FIGURE 6A

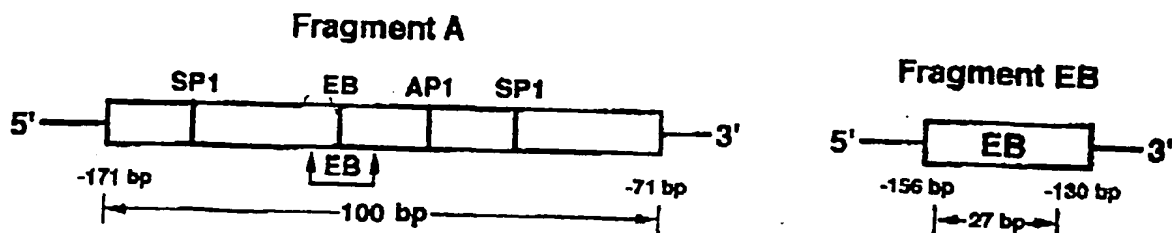


FIGURE 6B

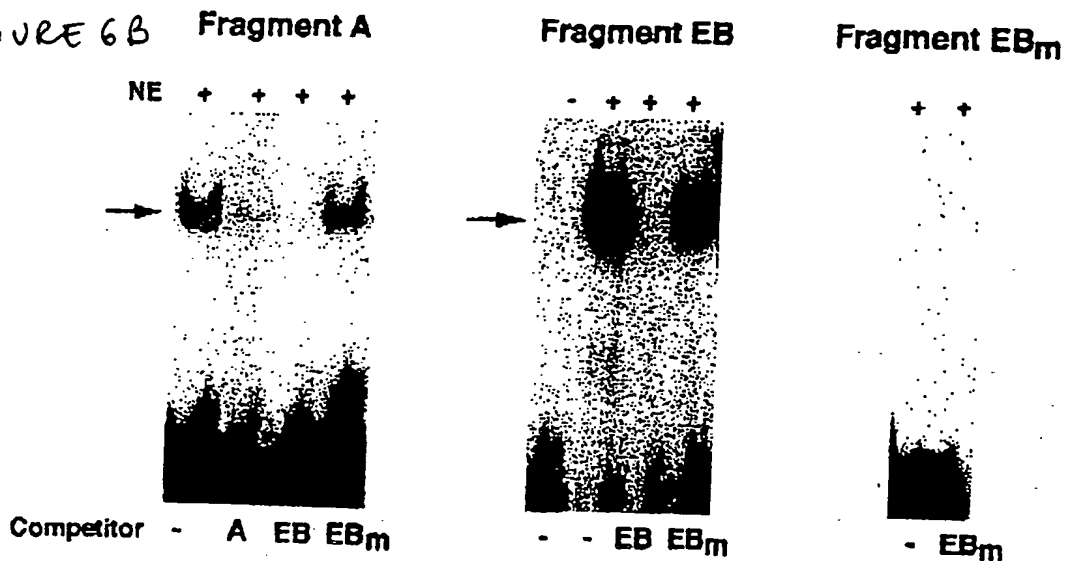


FIGURE 6C

